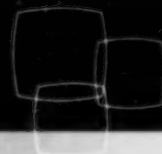




Industry
Canada

Industrie
Canada

WORKING PAPER SERIES / COLLECTION DOCUMENTS DE TRAVAIL



**PRIVATE AND SOCIAL RATES OF RETURNS
TO R&D IN OECD COUNTRIES:
HOW DIFFERENT ARE THEY ACROSS INDUSTRIES?**

Ram C. Acharya, Industry Canada

Working Paper 2008-04

The Economic Research and Policy Analysis Branch's Working Paper Series is intended as a means of diffusing research findings from Industry Canada projects for discussion and comment.

Working Papers are circulated in the language in which they were written. The papers reflect the views of the authors and no responsibility for them should be attributed to Industry Canada or the federal government. Comments on the papers are invited and may be sent directly to the authors.

La série Documents de travail de la Direction générale de la recherche économique et de l'analyse des politiques se veut un moyen de diffuser les résultats des recherches issues des projets d'Industrie Canada dans le but de favoriser la discussion et la réception d'observations.

Les documents de travail sont diffusés dans la langue dans laquelle ils ont été écrits. Les opinions qui y sont exprimées sont celles des auteurs et n'engagent pas Industrie Canada ou le gouvernement fédéral. Le lecteur est prié de faire part de ses commentaires aux auteurs.

Working Paper Series / Collection Documents de travail

PRIVATE AND SOCIAL RATES OF RETURNS TO R&D IN OECD COUNTRIES: HOW DIFFERENT ARE THEY ACROSS INDUSTRIES?

Ram C. Acharya, Industry Canada

Working Paper 2008-04

IC 60494

To obtain copies of documents published under the
Working Paper Series, please visit:
[http://www.ic.gc.ca/epic/site/eas-
aes.nsf/en/h_ra01967c.html](http://www.ic.gc.ca/epic/site/eas-aes.nsf/en/h_ra01967c.html) or contact:

Publication Coordinator
Economic Research and Policy Analysis
Industry Canada
10th Floor, East Tower
235 Queen Street
Ottawa, Ontario K1A 0H5

Tel.: (613) 952-6411; Fax.: (613) 991-1261
E-mail: mepa.apmc@ic.gc.ca

Pour obtenir des exemplaires des documents publiés
dans la collection des documents de travail, cliquer
sur : [http://www.ic.gc.ca/epic/site/eas-
aes.nsf/fr/h_ra01967f.html](http://www.ic.gc.ca/epic/site/eas-aes.nsf/fr/h_ra01967f.html) ou s'adresser à :

Coordinatrice des publications
Recherche économique et analyse des politiques
Industrie Canada
10^e étage, tour Est
235, rue Queen
Ottawa (Ontario) K1A 0H5

Tél.: (613) 952-6411; Fax.: (613) 991-1261
Courriel: mepa.apmc@ic.gc.ca

Acknowledgements

I am grateful to Jeffery Bernstein, Barbara Fraumeni, Surendra Gera and Sean Rogers for providing helpful comments. I also thank seminar audiences at Industry Canada, the 2007 Ottawa Productivity Conference, the 2007 Canadian Economics Association's meeting, and the 2007 NBER Summer Institute for comments. The views expressed are those of the author and not those of Industry Canada or the Government of Canada.

Author contact information: Ram Acharya, Industry Canada, 10 East – 235 Queen Street
Ottawa, Ontario K1A 0H5, Email: acharya.ram@ic.gc.ca.

Abstract

Using data for 17 Organisation for Economic Co-operation and Development (OECD) countries over 30 years for 28 industries, the paper estimates private and social returns to research and development (R&D). Results show that the excess private return (excess over normal return to factor of production) that an R&D dollar in an industry generates to itself is positive for all industries ranging from 3 to 57%. R&D of some industries generates spillovers to others, making private returns different from social returns. The paper provides partial support to Griliches' conjecture as it finds that, in some cases, an idea generates a greater impact on other industries than where it is generated.

Key words: R&D, private return, social return, industry technology flow matrix

Résumé

En utilisant les données relatives à 28 industries de 17 pays de l'Organisation de coopération et de développement économiques (OCDE) sur 30 ans, l'étude estime les taux de rendement privé et social de la R-D. Les résultats montrent que l'excédent de rendement (le rendement qui dépasse le taux de rendement normal des facteurs de production) de un dollar de R-D dans une industrie est positif pour toutes les industries. Il va de 3 p. 100 à 57 p. 100. De plus, la R-D de certaines industries a des répercussions positives sur d'autres, ce qui cause la différence entre le taux de rendement privé et le taux de rendement social. L'étude corrobore partiellement l'hypothèse de Griliches selon laquelle, dans certain cas, une idée a plus d'effets sur une autre industrie que sur celle où elle a pris naissance.

Mots clés : R-D, rendement privé, rendement social, matrice de courant de technologie industrielle

I. Introduction

R&D is different from other inputs in a sense that the amount of R&D spent in one sector can have spillover impacts in other sectors of the domestic economy (domestic spillovers) and in foreign countries (foreign spillover). The existence of the spillover creates a wedge between private R&D return (return to itself) and social return (private return *plus* R&D benefits transmitted to others). Furthermore, R&D, an input for technological change, is considered to be an important determinant of differences in per capita income among countries, differences in the size of industries and differences in growth trajectories across time. Hence knowing private and social rates of return to R&D is important for both academia and policymakers. For quite a few years, economists have been trying to quantify the amount, the direction, and the channels of R&D effect using a production function or, its dual, a cost function (for detail literature review on R&D spillovers, see Griliches (1992) and Nadiri (1993)).²

A common feature of the studies using a production function has been to estimate R&D spillovers, rather than social rates of return, by relating an industry's total factor productivity (TFP) to its own R&D and R&D of all other industries aggregated into one series. By construct, these models estimate private return to own R&D and the spillovers that this industry receives from other industries' R&D. The latter part is not useful in estimating social rate of return; for that, one needs to estimate the impacts of an industry's R&D on its own productivity and on other industries' productivity, not the impact of other industries' R&D on its productivity.

² Some of the studies in this context are Scherer (1982), Griliches and Lichtenberg (1984), Jaffe (1986) and Keller (2002a) which use a production function approach, and Bernstein and Nadiri (1988) which use a cost function approach. Bresnahan (1986) has developed a different methodology, which could be used to estimate the impact of R&D spillovers from an upstream industry to its downstream ones. Keller (2004) summarizes the methodology used in estimating foreign R&D spillovers which is similar in nature to estimating domestic spillovers.

However, there is a tendency to describe the *sum* of private return to own R&D and the spillovers that the industry receives from other industries' R&D as social rate of returns.³ This misnomer has happened despite the fact that Griliches (1992) appropriately named the former as "within" impact and the latter as impact "from outside", without implicating the sum of two as social return.

Furthermore, we also have methodological concerns on estimation techniques used in these studies, especially on how other industries' R&D is aggregated into one series to obtain R&D impact from "outside". Correctly realizing that not all industries' R&D is equally spillable to a particular industry (industry heterogeneity in source of spillovers) but not being able to allow each industry's R&D as a separate variable due to data constraints, these studies aggregate R&D of different industries using weights. The weights are based on either input-output (I-O) coefficients as in Scherer (1982) and Griliches and Lichtenberg (1984), or technology flow matrices as in Jaffe (1986). We argue that this practice might lead to estimation bias that is avoidable using the rich dataset and methodology used in this paper.

The I-O coefficient method of aggregation assumes that the R&D spillover of industry j to industry i is proportional to industry i 's input purchase from industry j . This is tantamount to assuming that R&D spillovers are siphoned mostly through the use of intermediate inputs and that too proportionately. If R&D spillovers were mainly of public nature so that any industry can grab it, then the I-O based weighting method would generate biased estimates.⁴ Similarly, a convincing argument can be made against using technology flow matrix as a weighting scheme.⁵

³ Nadiri (1993), has named them as "direct" effect and "indirect" effects, respectively and in a note underneath Table 2b, he provides a formula for social rate of return as the *sum* of these direct and indirect rates of returns. Although that is true for some studies he quotes, especially those that are based on a cost function approach, it is not the case for studies that are based on a production function approach. Jones and William (1996) in Table 1 imply the same.

⁴ Furthermore, as Griliches (1979) has succinctly described, the "embodied" nature of spillover could be more of a rent transfer rather than knowledge transfer. It could be a rent because R&D intensive inputs are purchased from

The other methodological issue with these studies is that they have not introduced heterogeneity in own-industry output elasticity to R&D. Rather they assume a common elasticity for all industries, yet another potential source of estimation bias.

To recap, as far as we are aware, there are no studies estimating both private and social rates of return to R&D using production function approach. And those that estimate R&D spillovers using this approach have methodological issues to deal with regarding how to aggregate R&D of other industries and how to incorporate heterogeneity in estimating own return. To address these gaps, this paper estimates the private and social rates of return to R&D by using data for 17 OECD countries for 28 manufacturing and services industries over 30 years (1973-2002). We do that by estimating heterogeneity in inter-industry spillover (rather than assuming based on I-O coefficients or technology flow matrix) and also by allowing heterogeneity in own R&D return.

Question arises, why not use cost function approach and estimate private and social rates to returns as is done in Bernstein and Nadiri (1988). Since there are no input prices including wage rates and rental rates at the level of detail required, we will not be able to use a cost function approach. Besides these data issues, as the coefficients to be estimated increase quite fast in cost function estimations as more industries are added, it cannot be used for very detailed industry structures, which we think is essential for estimating return to R&D.⁶ Looking at the

other industries at less than their full "quality" price. In that case, spillovers are just consequences of conventional measurement problems (not measuring all input prices correctly), not really a case of pure knowledge spillovers. If that is the case, then weighting R&D by I-O coefficients in measuring knowledge transfer may not be appropriate.

⁶ This methodology is based on defining "closeness" of industries by the nature of patents they have and assuming that the R&D spillovers are proportional to the measure of "closeness" matrices. Leaving aside the very vexing questions of what is "close", let us take a simple scenario that suffices to make our point. In a world where a conglomerate could engage on several distinct patents, it may not be appropriate to assume that the knowledge transfer is proportional to technology "closeness". As far as R&D spillovers are concerned, patents in two seemingly distinct industries generated by a conglomerate might be "closer" than patents in seemingly close industries but generated by rivals.

⁷ So far the studies based on a cost function approach have samples of only a few industries (five in Bernstein and Nadiri, 1988). Either because of data constraints or because of methodological constraints, the researchers use only a

scope of the paper (including industry, country and time dimensions and the nature of the questions we want to answer), it will be insurmountable to cover it using a cost function approach.

The paper makes a number of contributions to the literature. First, the paper decomposes the total impact of business R&D of each of the 10 most-R&D-intensive and the biggest-R&D-spender industries (which contributes about 80 percent of global R&D capital) into own-industry effect (impact it generates to its own TFP—spillover to its own TFP) and inter-industry effects (impact on all of the other remaining 27 industries' TFP) by allowing R&D of those industries to enter as separate sources of R&D spillovers. With these results, the paper is able to rank industries in terms of both private and social returns to R&D.

Second, by doing so, the paper allows heterogeneity in both own-industry and inter-industry effects of R&D across industries. Besides tackling the issue of whether there is a variation across industries in the returns to R&D that they generate for themselves, the paper also deals with whether the benefits that a dollar of R&D in different industries spills to a given industry are different. These two phenomena can be regarded as source heterogeneity. The paper also deals with destination heterogeneity of R&D effect by developing industry-by-industry technology diffusion flow matrix, and showing how much spillover each industry emanates to others (where do the spillover destines).

Third, the paper tests if there are additional benefits to be obtained by input-using (downstream) industries from R&D of an input-producing (upstream) industry because of their production linkages. This will also be a test of whether the usual practice of representing the R&D spillovers heterogeneity by I-O coefficients could resemble the reality.

few industries leaving others out of the sample. They have not been able to look at the impact of these sample industries' R&D on those excluded industries. To the extent that these sample industries' R&D affects excluded

Fourth, this is the most comprehensive study of its kind, not only in terms of industry, country and time coverage, but also in terms of econometric treatment. To our knowledge, this is the first study developing detailed industry interdependence of technology diffusion at the OECD level. In terms of econometrics, we use both instrumental variable and control-function approaches to correct for endogeneity issue in production function.

Fifth, this is the only study that incorporates foreign R&D in estimating private and social rates of return to domestic R&D at such a level of industry disaggregation. Since foreign R&D is found to be a very important source of domestic TFP growth, exclusion of that variable might lead to specification bias.

The main results of the paper are that there are substantial own-industry positive effects of R&D, and their magnitudes vary immensely across industries from as low as 3% to 57%. The estimated heterogeneous own-industry effects across industries suggest that the usual way of estimating common coefficients on own-industry R&D effect for all industries might lead to biased estimates. The results also show that the private rate of return and R&D intensity does not seem to have any relationship, as more intensive industries do not necessarily have higher returns.

There is also heterogeneity in inter-industry spillovers. Looking at the aggregate OECD economy-wide impact, inter-industry R&D spillovers are present only in four industries. The R&D of pharmaceutical industry is estimated to have a 69% inter-industry spillovers rate of return, whereas the chemical (excluding pharmaceutical) industry has 33%, the motor vehicle industry has 4% and the radio, TV and communication equipment industry has 2%.

industries' cost functions, the estimates might be biased.

Furthermore, there is heterogeneity on what each industry spills to individual 28 industries and also what each of these 28 industries receive from R&D of a given industry. For example, a 10% increase in the R&D of pharmaceutical industry increases TFP of aircraft and spacecraft industry by 2.8%, that of office accounting and computing machinery industry by 1.2% and that of electrical machinery industry by 1.4%. On other aspect, the TFP of machinery and equipment industry rises by 1.6%, 1.1% and 0.5% with the 10% R&D capital increase in pharmaceutical, chemical (excluding pharmaceutical) and motor vehicle industries respectively. These results confirm that there is a lot of heterogeneity both in the sources and in the destinations of R&D spillovers.

Results show that the transmission of inter-industry technology diffusion between upstream and downstream industries is positive but very small. On average, input-output linkages contribute less than 10% of total spillover. Hence, the practice of weighting other industries R&D spillovers in proportion to I-O coefficient is misspecified.

Following Griliches and Lichtenberg (1984) the *sum* of own and inter-industry effects is called social net excess rates of return (SNERR).⁷ If we make a reasonable assumption that the remuneration paid to labor and capital (normal private return) is the same across industries (the assumption of perfect factor mobility across industries), then the *sum* of this common return and SNERR yields social rate of return, and in this case, own-return can be regarded as private return and the SNERR can be regarded as social return. Viewed in this way, for six industries social return is equal to private return, as they do not generate inter-industry effect. Among those four industries that generate inter-industry effects, the social rate of return is almost four times its

⁷ It is a *social* rate of return because it is based in output rather than profit calculations. It is *net* because we have used the depreciation (rate of 15%) while computing R&D capital from investment. It is *excess* because the conventional inputs of labor and capital already include most of the R&D expenditures once at "normal" factor prices. See Schankerman (1981) for discussion on why it cannot be named exactly as "excess return".

private return for pharmaceutical industries and one and half times its private return for the motor vehicle industry. For the remaining two, since the inter-industry effect is very small, private and social rates of return do not vary much.

We examine the conjecture that Griliches made, "it is quite possible for an idea to have its entire effect elsewhere than where it was originated (1992, pp. S40-S41)." Our results only partially support his conjecture as not all benefits are spilled to other industries (each industry has positive own-industry R&D effect), but in some cases, it is true that the idea generated in an industry has a larger impact on other industries than on itself.

That much is for the impact of domestic R&D. On foreign R&D, interestingly it has a very strong and robust positive intra-industry effect, but there is no inter-industry foreign R&D effect.

The rest of the paper will proceed as follows. In section II we will describe the new dataset that is used for our empirical analysis, before turning to the estimation issues in section III. We provide the first set of empirical results in section IV. The rates of returns based on these empirical estimations are presented in section V. In Section VI we provide another set of empirical results, an industry-by-industry technology flow matrix. Section VII concludes the paper.

II. Data

This is the most detailed dataset so far used to measure the rate of return to R&D which has data on 17 countries and 30 (1973-2002) years. The dataset comprises 28 industries, with 22 manufacturing; four services; the electricity, water and oil industry, and the construction industry. These industries are in International Standard Industrial Classification (ISIC), revision 3 codes. Details on ISIC codes and industry names are provided in Table A1 (Appendix A). Details on data sources, construction and estimation are provided in Appendix B. In summary, we have compiled data on value added, employment, investment, bilateral trade, input-output coefficient, and R&D expenditures for 17 countries, and 28 industries for a period of 30 years.

Internationally comparable figures on employment, output, and sectoral prices come from the OECD's Structural Analysis (STAN) database for 1973-78 and from the Groningen Growth and Development Centre database for the years 1979-2002. Also from the OECD's STAN database comes data on investment. Data on business R&D are from the OECD's Analytical Business Expenditure on Research and Development (ANBERD) database, and data on the bilateral trade are from the OECD's Bilateral Trade (BTD) database. Investment and R&D expenditure data for each industry in each country and year, which were converted to US\$ PPP at 1995 price from national currencies in current price, are used to calculate physical capital and R&D capital based on a perpetual inventory method.

The data are compiled in such a way that labor and capital inputs also include the R&D expenditures, and there is no way to separate them from total labor and capital costs. Hence, the return to R&D estimated using these data would be an excess return (over normal return paid to conventional labor and capital input), a spillover to TFP, so to speak. The number of workers

measures labor inputs. The measure of output in this analysis is value added, since internationally comparable data on intermediate inputs are not available.

Since availability of data varies by individual data series, our sample is an unbalanced panel. With 17 countries and 30 years of data, there is a maximum of 510 observations for each industry. For each country, there are 840 possible observations. As the lower part of Table A1 shows, data availability by industry varies somewhat; especially the R&D data on services industries (#23 through #26) which are missing more. The dataset is complete for many countries. The major exceptions are (i) Belgium, for which R&D data becomes available only in 1987; (ii) Ireland, for which investment data starts only in 1992, and (iii) South Korea, where R&D data is only recorded from 1995 onwards. In addition, there are some missing values during the 1970s. However, we do not expect this missing data to have an important influence on the results given the wide spread of data across countries, industries and years.

Table A2 provides information on industry's R&D intensity (R&D over value added) by country (the industries are sorted by R&D intensity during 1973-2002 in ascending order). Data show that the average R&D intensity varies a lot by sector; also for a given industry, it varies a lot by country. OECD has labeled first five most-R&D-intensive industries as high-tech, and the next five industries (from #6 to #10) as medium-high-tech. The next six (from #11 to #16) are labeled as medium-low-tech, followed by another six (from #17 to #22) as low-tech industries. Then there are four services industries (#23 through #26) along with the remaining electricity and construction industries.

A summary account based on Tables A2 and A3 along with share of value added of the most-R&D-intensive 10 industries, 3 sectors and two industries is presented in Table 1. As far as estimating the rate of returns to R&D is concerned, our focus will mainly be on 10 high- and

medium-high-tech industries. These 10 industries are also the biggest-R&D-spenders with the exception for railroad and transportation equipment (Table A2). Together they account for about 77% of the OECD countries' R&D capital; other 18 industries contribute only the remaining 23%. In the appendix, Tables A4-A6 provide summary statistics on employment, capital stocks, and R&D stocks by industry and by country.

Table 1. R&D intensity and shares of R&D and value added in OECD, 1973-2002

Industry Name	ISIC	R&D intensity	R&D share	Value added share
High- & medium-high-tech manufacturing		14.2	76.7	9.2
1 Aircraft and spacecraft	353	38.3	9.8	0.4
2 Office, accounting and computing machinery	30	30.5	6.8	0.4
3 Radio, TV and communication equipment	32	26.8	16.0	1.0
4 Pharmaceuticals	2423	21.2	7.2	0.6
5 Medical, precision and optical instruments	33	14.4	5.8	0.8
6 Motor vehicles, trailers and semi-trailers	34	13.3	11.9	1.4
7 Electrical machinery and apparatus, nec	31	8.9	5.0	1.0
8 Chemicals (excluding pharmaceuticals)	24ex2423	8.9	7.9	1.5
9 Railroad and transport equipment nec	352+359	8.1	0.5	0.1
10 Machinery and equipment, nec	29	4.8	5.8	2.1
Medium-low-tech manufacturing	23, 25, 26, 217, 272, 351	3.0	6.1	3.5
Low-tech manufacturing	15-16, 17-19, 20, 21-22, 28, 36-37	1.1	5.2	8.3
Services	60-63, 64, 71-74, other services	0.2	8.6	69.5
Electricity and Construction	40+41, 45	0.6	3.0	9.5
Average/total	Total of 28 industries	1.7	100	100

Other services include ISIC 50-52, 55, 65-67, 70, 75-99. The data on all these ISIC are not available separately but comes in the aggregate.
nec means "not elsewhere classified".

III. Estimation

Consider the Cobb-Douglas production function for industry i at time t in country c :

$$(1) \quad Y_{cit} = A_{cit} K_{cit}^{\beta_k} L_{cit}^{\beta_l},$$

where $i = 1, \dots, 28$; $c = 1, \dots, 17$; and $t = 1973, \dots, 2002$; Y is value added; K is capital, L is labor, and β_k and β_l are the elasticities of capital and labor, respectively. The term A in equation (1) is an index of technology, or productivity, which is treated as a residual contributor. It follows that

$$(1') \quad y_{cit} = \beta_k k_{cit} + \beta_l l_{cit} + a_{cit}$$

where for any variable Z , $z = \ln Z$. Hence

$$(2) \quad a_{cit} = y_{cit} - \beta_k k_{cit} - \beta_l l_{cit}$$

There are several factors that could affect technology variable a . Let us define those arrays of variables by \mathbf{X} . Then

$$(2'') \quad a_{cit} = \beta_0 + \mathbf{X}\mathbf{B} + \varepsilon_{cit}$$

Replacing a_{cit} in (1') by expression in (2''), (1') yields the main estimation equation

$$(3) \quad y_{cit} = \beta_0 + \beta_k k_{cit} + \beta_l l_{cit} + \mathbf{X}\mathbf{B} + \varepsilon_{cit}$$

The estimated value of the error term, ε_{cit} , is total factor productivity, or a_{cit} . A number of generic issues exist in the estimation of the capital and labor coefficients. In the multivariate regression context any bias in β_k and β_l generally leads to biases in \mathbf{B} as well. A major econometric issue confronting production function estimation is the possibility of simultaneity (endogeneity), a problem referring to the fact that at least a part of TFP (which basically is a error term) will be potentially observed or predicted by firms when they make input decisions. If that is the case, then firms' optimal choice of inputs l_{it} and k_{it} will generally be correlated with the observed or predictable productivity shock (endogeneity issue). Specifically, one can split the error term ε_{cit} into two elements

$$(4) \quad \varepsilon_{cit} = \omega_{cit} + u_{cit}$$

where ω_{cit} (which could be a determinant of productivity) is observed by agents who choose the inputs early enough to influence decisions, while u_{cit} is the true error term that may contain both unobserved shocks and measurement errors (in y_{cit}). This implies that OLS will generally not yield unbiased parameter estimates because $E[l_{cit} \varepsilon_{cit}] \neq 0$ or $E[k_{cit} \varepsilon_{cit}] \neq 0$, or both. Even if ω_{cit}

is time-invariant across groups (country-by-industry combination) so that $\omega_{it} = \omega_{ci}$ or group-invariant across time $\omega_{it} = \omega_t$ (shocks affecting all industries in the sample), the problem still remains. We will employ several estimators to address this issue. First, we assume that the unobserved term ω_{it} is given by country-, industry-, and time-effects that are fixed and can be estimated as parameters:

$$(4') \quad \varepsilon_{it} = \eta_c + \mu_i + \tau_t + u_{it},$$

If (4') holds, OLS will yield consistent and unbiased estimates. However, there are two drawbacks of this method. First, a fixed effect estimator uses only the across time variation, which tends to be much lower than the cross-section one, thereby identifying coefficients weakly. Second, the assumption that the term can be represented by fixed effects may not always be reasonable, making the whole procedure invalid.

Second, we will employ the General Method of Moments (GMM) techniques developed by Arellano, Blundell, Bond, and others (Arellano and Bond 1991, Blundell and Bond 2000), which is suitable even in the presence of endogenous regressors, which may be the case under the simultaneity problem. Assume that

$$(4'') \quad \varepsilon_{it} = \zeta_{it} + \tau_t + u_{it},$$

where year fixed effects (τ_t) control for common macro effects; ζ_{it} is the unobservable industry component, and u_{it} is a productivity shock following an AR(1) process, $u_{it} = \rho u_{it-1} + \eta_{it}$. Then the firms' input choices are endogenous with respect to ψ_{it} . Assumptions over the initial conditions and over the serial correlation of u_{it} yield moment conditions for combining equations in levels (of variables) with equations in differences (of variables) for a System GMM approach which uses lagged values to construct instrumental variables for current variables.

Third, we adopt the approach developed by Olley and Pakes (1996), who address the simultaneity problem by using firm's investment (observed input) decision to proxy unobserved productivity shock, ω_{it} .⁸ A standard approach of Olley and Pakes (OP) is that contemporaneous values of capital and investment are orthogonal, and a higher value of ω_{it} induces a higher investment today. Hence, in their model, capital is treated as predetermined and labor is the only variable factor. They maintain that the investment function which is an increasing function of ω_{it} can be inverted to obtain $\omega_{it} = f^{-1}(i_{it}, k_{it})$. Next, this expression is substituted in the production function to control for ω_{it} and the labor coefficient is estimated at the first stage. Once a consistent estimate of ω_{it} is obtained, the source of the potential endogeneity problem is eliminated, and in the second stage, the capital coefficient is estimated.⁹ We will employ both a variant of Olley and Pakes' two-step procedure as well as the more recent one-step GMM procedure proposed by Wooldridge (2005). Wooldridge maintains that two-step OP estimators are necessarily inefficient because it ignores contemporaneous correlation in the errors across two equations and they do not efficiently account for serial correlation or heteroskedasticity. By contrast, he states that the GMM approach uses the cross-equation correlation to enhance efficiency, as well as an optimal weighting matrix to account for serial correlation and heteroskedasticity. Another benefit of the GMM approach is that any over-identification assumptions implied by economic theory are easily tested without bootstrapping.

⁸ The idea is that period t capital stock is determined at period $t-1$. Hence the productivity shocks between period $t-1$ and t do not determine the capital stock in period t and hence investment. However, labor, which is more variable and is decided at period t , is potentially correlated with the innovation component, ω_{it} , of error term.

⁹ See also Griliches and Mairesse (1998) and Akerberg, Caves, and Frazer (2005) for a discussion of these assumptions.

IV. Empirical Results

We will expand X in Equation (3) to estimate the impacts of own-industry and inter-industry domestic R&D and intra-industry and inter-industry foreign R&D on productivity. We start estimation using all four OLS, OP, IV System GMM and OP Wooldridge (OPW) methods based on different assumptions on the regression error ε_{cit} . At different stages of the estimation, we abandon using OP and IV System GMM and stick with OLS and OPW. For point estimation of rates of return to R&D, our preferred estimation technique is OPW, whereas OLS estimates are used to compute range of returns.

1. The role of domestic and foreign R&D: the basics

We start with a simple version of equation (3) where technology a of industry i is determined by domestic R&D of industry i . Furthermore, as foreign R&D is found as an important promoter of domestic technology, we consider it as another determinant. The assumption here is that R&D of one industry does not affect the productivity of other industries—neither at home nor from abroad—i.e. no inter-industry R&D effect. Specifically, we estimate the following equation:

$$(5) \quad y_{cit} = \beta_0 + \beta_k k_{cit} + \beta_l l_{cit} + \gamma r_{cit} + \phi^A \sum_{c' \neq c, c' \neq 0} m_{c'it} r_{c'it} + \varepsilon_{cit}$$

In this expression, r_{cit} is the own-industry domestic R&D of country c in industry i at time t , and hence γ measures the output elasticity with respect to domestic own-industry R&D.¹⁰ Here $r_{c'it}$ is the amount of foreign R&D stock from country c' in industry i in year t . The foreign R&D stock is defined as in Coe and Helpman (1995) such that the R&D stock for country c for industry i from country c' , is obtained by the product of the share of country c' in total imports of country c

in industry i from 17 sample countries and the amount of R&D stock in country c' in industry i' .

For a given country and year, we compute the share of a country c' in total imports of c as

$m_{cc'u} \equiv M_{cc'u} / \sum_{c' \in C' \neq c} M_{cc'u}$ where $M_{cc'u}$ is the import of country c from country c' and hence

$\sum_{c' \in C'} m_{cc'u} = 1$. Since the focus of the paper is not to estimate the foreign R&D spillover by

country, all other 16 countries' R&D enters as an aggregate explanatory variable with a single coefficient ϕ^i . The super-script A is used to indicate "intra"-industry effect—the effect of R&D of foreign industry i on productivity of same industry at home.

The results are presented in Table A7. Since this is just a starting point of empirical results, we comment on them only generally. The first column introduces only labor and capital in the equation, whereas the second column introduces own-industry domestic R&D, and the third one adds intra-industry foreign R&D as well. The fourth column is like column 3 except for using country and industry as two separate fixed effects, it uses the interaction of industry and country as fixed effects.¹¹ The estimators in column 4 are called *within* estimators. The coefficient of labor, which is 0.645 in the first column, declines substantially in the second one when domestic R&D is incorporated and rises slightly when foreign R&D is also introduced (column 3). In within-estimation, however, β_l is much lower, and β_k is much higher.

Since OLS may suffer from endogeneity problems, in other columns, we compare the least squares estimates with alternative estimators. Column 5 employs the IV System GMM estimator (Blundell and Bond, 2000). Labor, capital and own-industry R&D are treated as endogenous and are instrumented by l_{t-2} , l_{t-3} , k_{t-2} , r_{t-2} . Given three endogenous variables

¹⁰ We have computed physical capital stocks using the perpetual inventory method and depreciation rate of 5%. The R&D stocks are computed using a rate of depreciation of 15%. The labor measure is the total number of employees.

¹¹ Column 3 has $C + 1 = 17 + 28 = 45$ fixed effects, whereas column 4 has $C \times I = 17 \times 28 = 476$ fixed effects.

(l_t, k_t, r_t) , there is one overidentifying restriction, and the p -value of 0.720 for the Sargan test of overidentification statistic states that we cannot reject the null hypothesis that the instruments, as a set, are exogenous. Besides as desired by the model, the AR(1) test at 5% level rejects the null hypothesis that there is no first-order autocorrelation, and AR(2) test cannot reject the null that there is no second-order autocorrelation at a reasonable level of significance.

Specification (6) is based on the OP original two-step procedure. In step one, the unobservable $e_{i,t}$ (see equation 4) is approximated by a third-order polynomial in investment and capital, which allows the identification of β_l . In the second step, the assumption that capital is uncorrelated with the innovation, $e_{i,t}$, which follows a first-order Markov process, ensures the identification of β_k . In column 7, we show the results of implementing the OP estimator in the one-step GMM procedure recently proposed by Wooldridge (2005).¹²

In our estimation, β_l ranges from 0.358 to 0.645, and β_k ranges from 0.128 to 0.415. One would be able to check which estimation method generates more realistic elasticities if the data had supported constant returns to scale market structure, in which case, the labor elasticity would have been equal to labor share in value added. However, here in all specifications, the null hypothesis of constant return to scale of labor and capital inputs is rejected (p -value of 0.00). Hence at this stage, we have no way of knowing estimators based on which methods are better.¹³ Nevertheless, since β_k is unusually high in *within* estimator and insignificant in an OP approach, we will not use these two methods of estimations for the remaining part of the paper.

¹² This assumes that $e_{i,t}$ is a random walk (not only first-order Markov), and the identification for both β_l and β_k comes solely from moment conditions that correspond to Olley and Pakes' second stage.

¹³ In any case, the average labor share in the data, which is computed as the average of labor compensation over value added, is 0.62. The median is 0.65.

In the estimation, the domestic R&D elasticity ranges from 0.15 to 0.18. With average value added of PPP \$31323 and average domestic R&D capital of PPP \$2883, the rates of return are in the range of 163% to 195%.¹⁴ At such an aggregate level, it is not unusual to have such a high rates of return to R&D.

2. Role of domestic R&D: inter-industry effects

In the above estimation, we entered the domestic R&D in a way that the R&D of one industry affects only its own productivity without any spillover to other industries. We also forced the output elasticity with respect to own-industry domestic R&D, r , to be the same for all years, countries and industries. These are rather strong assumptions, and we relax both of them in this section. First, we allow each of the 10 most-R&D-intensive industries (given in Table 1) to have R&D spillover impacts on all the other 27 industries (inter-industry effect) but still maintain the assumption that the own-industry R&D effect of all 28 industries is the same. The estimating equation is:

$$(6) \quad y_{c,t} = \beta_0 + \beta_k k_{c,t} + \beta_l l_{c,t} + \gamma r_{c,t} + \sum_{i' \in \{1, \dots, 10\}} \gamma^{i'} r_{c,t}^{i'} + \phi^{i'} \sum_{i' \in \{1, \dots, 10\}} m_{c,i'} r_{c,t} + \varepsilon_{c,t} \quad i \in 28.$$

Equation (6) is like Equation (5) except that the inter-industry R&D component, $r_{c,t}^{i'}$, is added in the former. As before, $r_{c,t}$ is the series of own-industry R&D of all 28 industries, and $r_{c,t}^{i'}$ is R&D of ten industries as ten separate series. We have super-scripted the coefficient and r by i' to indicate "inter"-industry entries. Coefficient γ measures own-industry R&D effect, whereas each of ten $\gamma^{i'}$ coefficient measures inter-industry R&D spillovers of i' (each of ten) industry. To avoid double accounting, under the series $r_{c,t}^{i'}$, the R&D values for country c , industry i' and time

¹⁴ The lower range rate of return for domestic R&D will be 1.63 (= 0.15*31323 / 2883).

i are set to zeros as these values already enter under series r_{cit} when i' is equal to i . Hence, in series r_{cit}' , R&D values of industry i' are repeated for all 27 industries except for itself.

The results are given in the first three columns of Table A8, which provide two lessons. First, if R&D is not allowed to generate inter-industry spillover, then its own-industry impact is estimated with upward bias. The output elasticity of domestic own-industry R&D, which is estimated in the range of 14 to 16%, is about 2 percentage points lower than those reported in Table A7, irrespective of whichever estimation technique is used. Second, own-industry effect is a lot higher than economy-wide inter-industry effect. The common own-industry effect ranges from 14 to 16% at the upper part of the Table 8, whereas the inter-industry effects at the lower part of the table are less than that for all industries and specifications.

Knowing that if inter-industry effects are not allowed to play role, own-industry effect will be overestimated, should we be including remaining 18 (28 minus 10) industries' R&D also as separate inter-industry R&D spillers in the model to have the coefficients unbiased? The inclusion of all 28 industries will be an impossible task given the number of parameters one has to estimate. So what we do next is to check own industry impact of the remaining 18 industries' R&D by separating them from the 10 most-R&D-intensive industries. If the own-industry effect of these 18 industries is small, then we could ignore their inter-industry effects (as the latter are known to be even smaller than the former from the result we just discussed), as they would not be creating omission bias. We estimate the following equation:

$$(7) \quad y_{cit} = \beta_0 + \beta_k k_{cit} + \beta_l l_{cit} + \gamma r_{cit}^{18} + \sum_{i' \in \{1, \dots, 18\}} \gamma_i r_{cit}^{i'} + \phi^A \sum_{i' \in \{1, \dots, 18\}} m_{cit}^{i'} r_{cit}^{i'} + \varepsilon_{cit} \quad i \in 28.$$

In the own-industry R&D column, r_{cit}^{18} , the R&D values of only 18 industries are entered and the R&D values of the ten industries are set to zeros, whereas in the vectors of ten $r_{cit}^{i'}$ series the data on R&D of each of these 10 industries are entered for all 28 industries including for

itself. The difference between Equation (6) and Equation (7) is that, in the former γ measured own-industry impact of all 28 industries, whereas in the latter it does that only for 18 industries. Furthermore, in Equation (6) each γ_i measured only the inter-industry R&D effects of each of the 10 industries, whereas in Equation (7), it measures the total (own- and inter-industry) R&D effects of each of them.

The results are provided in the last three columns of Table A8, and the following points can be made. First, our approach of considering only 10 industries as separate spillers is justified, as there is either no impact or a very small impact of the other 18 industries' R&D even in their own productivity. The own-R&D effect, in the first three columns, was generated by 10 industries. Once they are taken out, the output elasticity with respect to own R&D not only falls substantially (1 and 5% in columns 4 and 6 compared to the range of 14 to 16% in columns 1 through 3) but also is insignificant in one case (column 5). Second, the total effect of R&D varies substantially by industry. Therefore, while estimating rates of returns to R&D more disaggregate industry dimension is recommended to avoid aggregation bias. Third, the total effect (the sum of own and inter industry effects) of R&D could be negative, implying that either own-effect is negative or negative inter-industry effect outweighs the positive own-industry effect.

What about the negatively significant coefficients? How is one supposed to interpret these "wrong" signs? Generally, we do not expect the R&D of one industry to have a negative effect on itself and other industries. However, the negative TFP elasticity with respect to R&D is not new, but researchers have a tendency either not to highlight this or to ignore them saying the coefficients might be picking something else.¹¹ The endogenous growth model however,

¹¹ Zvi Griliches and Frank Lichtenberg (1984) using data on 27 industries from 1959 to 1976 estimate constant elasticity equation like ours find that the coefficient on R&D to level of TFP were negative in all six specifications, and insignificantly different from zero in all but one (Table 3). In another paper, using firm level business units data from 1971 to 1980, Clark and Griliches (1984) find that the coefficients of variable "mis", which is defined as ratio

recognizes that duplication and overlap externality in research, called "stepping on the toes", reduce the number of innovation produced by a given amount of research, thereby affecting productivity negatively (Jones and Williams, 1998).

To check whether the negative results are due to omitted variables, following the literature that market competition affects productivity, we also performed estimations by controlling for competition variables (proxied by share of imports in value added and mark up rate). The results did not change; they remained almost intact.¹⁶ Another possible suspect for "wrong" sign is multicollinearity as a result of simultaneous entries of 10 industries' R&D. There is no evidence of that either.¹⁷ So with this discussion, we are tempted to rely more on line of endogenous growth literature in a sense that there could be an occurrence of "stepping on the toes" in the innovation field. Since the structure of this paper is not suited for an estimating endogenous growth model, we cannot do justice in explaining the negative signs.

Using the results in Table 8, we want to discuss the merit of the three estimation techniques used. The bottom line is that the results seem to be more precisely estimated under OPW and OLS than under GMM. Even though in many cases, the magnitude of GMM coefficients are not that different from those under OPW, the latter are mostly insignificant because of their large size of standard errors. Even between OPW and OLS, the magnitudes of coefficient and standard errors suggest that the results are slightly more precise under OPW than under OLS. In view of this discussion, our preferred estimation technique is OPW, and we will use the results under this technique to estimate the rates of returns to R&D. We also continue to use OLS but drop the use of IV System GMM for the remaining part of the paper.

of product R&D expenses to total R&D expenses as an additional variable along with R&D intensity to sales, are negative in all cases. Keller (2002a) finds that the domestic industry's TFP is negatively affected by foreign countries' other industries' (not the same industry) R&D.

¹⁶ On competition and TFP, see Nickell, 1995 and Disney et al., 2003.

3. A conduit of R&D spillover: production linkage

So far, we have concentrated on the disembodied R&D benefits that are of a public good nature for any industry to capture from any other industries. It is quite likely that besides benefiting from a public good nature of other industries' R&D, industries might obtain additional benefit from the R&D of input-producing industries.¹⁸ To test this hypothesis, we introduce input use as a transmission mechanism for R&D spillovers from input-producing industry to input-using industry, a forward linkage phenomenon. This exercise is particularly important, as previous studies have used I-O coefficients as weights to aggregate other industries' R&D into one series, and this could be a check of whether the industry heterogeneity introduced that way reflects the reality. Our estimating equation is:

$$(8) \quad y_{i,t} = \beta_0 + \beta_1 k_{i,t} + \beta_2 l_{i,t} + \gamma_{i,t}^{18} + \sum_{j \in 10} \gamma_j r_{i,j,t} + \sum_{j \in 10} \mu_j n_{i,j,t} + \sum_{j \in 10} \theta_j n_{i,j,t} r_{i,j,t} + \phi^A \sum_{i' \in 28, i' \neq i} m_{i,i'} r_{i',t} + \varepsilon_{i,t}$$

There are two additional terms, $n_{i,j,t}$ and $n_{i,j,t} r_{i,j,t}$, from what we have in equation (7). The first term, $n_{i,j,t}$, is the share of industry j (each of ten most-R&D-intensive industries) in total input used by industry i from 28 industries, with the effect that $\sum_{j \in 28} n_{i,j,t} = 1$. For example,

$$n_{i,j,t} = N_{i,j,t} / \sum_{j \in 28} N_{i,j,t}, \text{ where } N_{i,j,t} \text{ is the output of industry } j \text{ that is used as input by industry } i.$$

Then for each of the 28 industries, we interact the input share of industry j to the R&D of

industry j , which is the second term, $n_{i,j,t} r_{i,j,t}$. Here, our interest is in parameters γ_j and θ_j ; γ_j

¹⁷ Even when we introduced only the R&D of industry 1, the most-R&D-intensive industry, as a single determinant, its coefficients were negative throughout all specifications.

¹⁸ The input-output table shows the input use by each industry from each of the sample industry. Since the input-output tables are available only for one year, there is no time dimension in this variable. Moreover, even if there were time dimensions, we may prefer to use fixed input-output coefficients, as we want to isolate the rent effect. The input-output matrices were available only for 14 sample countries and were missing for Belgium, Ireland, Korea and Sweden. The input-output tables we have used are for year 1995 for Australia, Finland, France, Germany and Spain. They are for 1998 for Great Britain and the Netherlands. For Canada, Denmark, Japan, Norway and the USA, the matrices are for 1997. For Italy, the input-output matrix used is for 1992. For four countries that have no input-

measures the disembodied (both own-and inter-industry) impact of R&D of industry i' on TFP of 28 industries, whereas $\theta_{i'}$ measures the spillovers of R&D in upstream industry to the productivity of downstream industry, the embodied impact. It means that each industry's output elasticity with respect to R&D of industry i' (one of the ten industries) is given by $\gamma_{i'} + \theta_{i'} \bar{n}_{i,j,i'}$, where $\bar{n}_{i,j,i'}$ is the mean of $n_{i,j,i'}$, i' industry's average share of its input used across 27 industries in total input use by these industries.

The results are presented in Table A9, using OPW for the first four columns and OLS for the last column. In the first column, the elasticity of industry 1's R&D is negative with coefficient of 0.022, whereas the indirect R&D effect through input use that this industry spills to other 27 industries is 0.139. Next, we introduce industries 2 and 3 in column 2 and 3 respectively. These first three columns are presented to see if the results in the last two columns are influenced by multicollinearity. It does not seem to be the case, as the coefficients of all three industries appear with the same qualitative feature even if the R&D of all 10 industries are introduced in the remaining two columns. In the last two columns, in addition to the R&D of all 10 industries, we also introduce their input shares, interaction terms of R&D and input share, own I.&D of domestic 18 industries and foreign intra-industry R&D (total of 32 control variables). However for brevity, the coefficient values of 10 industries' input share (μ_i) are not reported.

In qualitative terms, the direct impacts of R&D of all ten industries (R&D coefficients) are the same as in Table A8 (columns 4 and 6). In magnitudes, however, they are, as expected, somewhat smaller than those in Table A8 for some industries. The interaction terms are positively significant for all industries except for industry 9 in both specifications. There are no

output tables, we use other countries' tables as proxy. For Belgium, we use Germany's matrix, for Ireland we use

negative entries in the product terms in both specifications, so no risk of productivity loss in input-using industry with the rise in R&D in input-producing industry.

However, the R&D impacts through production linkages are very small, implying that the magnitude of innovation diffusion through an arm's-length transaction of input is minimal. As an example, let us take the case of four industries (3, 4, 6 and 8) whose direct R&D impact is positively significant. For industry 3, the input related coefficient is 0.143 which evaluated at the mean input share of 0.031 of industry 3 implies an indirect elasticity of 0.44% ($0.0044 = 0.143 \times 0.031$). For others, these elasticities are 0.31% for industries 4; 0.39% for industry 6, and 1.1% for industry 8.

As a result of their small magnitudes, the spillovers due to production linkage have negligible contribution to total R&D effect. For example, for industry 3, the production relation contributes only 10% ($=0.0044/(0.0044+0.039)$) of its total R&D effect. The entry 0.0044 is the product of the coefficient 0.143 and the input share of 0.031 and 0.039 is the coefficient of industry 3's R&D. In other words, out of total rate of return of 30% for industry 3's R&D through both sources (direct and indirect), only 3% comes from input linkages.¹⁹ Similarly, the rate of return to R&D from production linkages is 1% out of total of 29% for industry 4, 3% out of total of 19% for industry 6 and 10% for total of 89% for industry 8. Therefore, the role of input linkage in total R&D effect is either very negligible in magnitude as in industries 3, 4 and 6 or not more than 10% as in industry 8.

This result that the R&D of input-producing industries has no substantial impact on productivity of input-using industry is contrary to what is in the literature. Scherer (1982) found

that of Great Britain; for Korea, we use that of Japan and for Sweden, we use that of Finland.

¹⁹ This is obtained by multiplying the indirect coefficient of 0.0044 by average economy-wide value added to industry 3's R&D ratio of 6.9 ($=0.03$) and multiplying the direct coefficient of 0.039 by 6.9 ($=0.27$). The sum of two products is equal to 0.30 and the share of the first component is just 10%.

that R&D embodied in intermediate inputs (purchased from other industries), contributing perhaps even more than some of the R&D performed *within* the industry. Griliches and Lichtenberg (1984) also found uniformly larger estimates of the coefficients of embodied R&D than of the coefficient of own-product R&D, but not than those of own-process R&D. Our result also calls into question the methodology of aggregating other industries' R&D into one series using I-O coefficients, as these coefficients do not seem to have any proportional association with the amount of R&D benefits transferred from input-producing to input-using industries.

4. Decomposing R&D effects into own- and inter-industry effects

Before proceeding to decompose the total R&D effects into own- and inter-industry effects, let us recap some of the results that we have obtained so far: (1) to estimate the own-industry impact of R&D more precisely, we should allow avenues to inter-industry effects for the R&D of the 10 most-R&D-intensive industries, (2) however, we can ignore the inter-industry effect of the 18 less-R&D-intensive industries, (3) the returns to R&D (both own and inter) are heterogeneous across industries thereby warning us against aggregating them in a series, and (4) since the embodied impact of R&D from I-O linkages is very small, we can ignore it in the analysis. Not only are these results very important in themselves, more significantly, they provide us guidelines and safeguards on what we should incorporate in our estimating equation below in order to decompose the total (disembodied) R&D effect into own- and inter-industry effects for the 10 most-R&D-intensive industries. The estimating equation is:

$$(9) \quad y_{it} = \beta_0 + \beta_k k_{it} + \beta_l l_{it} + \gamma r_{it}^{18} + \sum_{i' \in 10} \gamma_{c'}^{i'} r_{c'i'}^{i'} + \sum_{i' \in 10} \gamma_{c'}^j r_{c'i'}^j \\ + \phi^4 \sum_{c' \in 17} m_{c'i'} r_{c'i'} + \phi^j \sum_{c'' \in 17} \sum_{i'' \in 28} m_{c''i''} r_{c''i''} + \varepsilon_{it} \quad i \in 28; i'' \in 28; i'' \neq i; i' = i; c' \neq c$$

Equation (9) is an extension of Equation (7), where domestic R&D coefficient γ_c in Equation (7) is decomposed into own-industry effect, γ_c^O (indicated by super-script *O*) and inter-industry effect, γ_c^I (indicated by super-script *I*).²⁰ Total effect of R&D is given by $\gamma_c^O + \gamma_c^I$. In foreign R&D, so far we were considering only intra-industry R&D (was the case in equation 7), here we add foreign inter-industry R&D effect, which is measured by the coefficient ϕ^I . In the above expression, $\sum_{i' \neq i} \sum_{c'} m_{c'i'} r_{c'i}$ is the foreign inter-industry R&D for industry *i*, where foreign R&D of all other 27 industries except foreign R&D from the same industry is aggregated. To obtain foreign inter-industry R&D for country *c* and industry *i*, we have added R&D across all countries except country *c* and across all industries except industry *i*.

Let us compare Equation (9) with equations used so far in the literature in estimating R&D spillovers using a production function approach. First of all, here, we introduce R&D of the same ten industries for both own and inter industry spillover R&D series, and as a result the total effect ($\gamma_c^O + \gamma_c^I$) is the return to the same R&D dollar. On the contrary, in the literature, for inter-industry R&D series, instead of using same industries' R&D, R&D of all other industries, other than that of industry *i* are used by aggregating it based on some sort of weight. As a result, the estimation of $\gamma_i^{O'}$ and γ_i^I will be the return for different industries' dollar. Furthermore, since all industries' R&D is aggregated in one series, the previous studies have only one γ_i^I coefficient. Furthermore, instead of estimating different $\gamma_i^{O'}$ coefficients for different industries (for example 10 in our case), the previous studies estimate only one common coefficient for all industries. And all these studies assume that $\phi^A = \phi^I = 0$.

²⁰ In series $r_{c'i}$, R&D values of each of the ten industries' enter for itself, with zero entries for all other remaining 27

Based on our findings above, we feel that the control variables introduced in the equation are sufficient. We do not expect to have specification bias while estimating Equation (9), at least not in terms of omitted variables, omitted channels of spillovers and aggregation. Yet, we have a reasonably elegant and comprehensive equation to estimate the private and social rates of returns to four-fifths of global R&D capital, which is the novelty of the paper. The results, which are reported in Table A10, are the key results of the paper. For ease of comparison and space saving, we report own-industry and inter-industry R&D effects in two separate columns even though they enter simultaneously as independent variables and the coefficients should have been reported in one column.²¹

Looking at the results, the own-industry R&D effects are positive for all industries in both specifications. Under OPW, the elasticity of output with respect to own R&D varies from as high as 0.327 for industry 2 (Column 1) to as low as 0.085 for industry 9. It means that if the R&D of industry 2 (office, accounting and computing machinery) rises by 10%, then its TFP level increases by 3.3%. Even the industry with the lowest elasticity, railroad equipment, will realize 0.9% increase in its TFP as a result of 10% increase in its R&D capital. Note that the own-industry elasticities of these ten industries are spread across the average own-industry elasticity of 0.143 in Table A8, column 3. In OLS, the highest elasticity is recorded for the same industry 2 at 0.407, whereas the lowest one is recorded for industry 10 at 0.083. Except for some industries, the magnitudes of elasticities are not much different between the two specifications. But the elasticities are higher under OLS than under OPW for all, except for industries 8 and 10.

industries. On the other hand, in series r_{it}^I , R&D values of each of ten industries enter for all other 27 industries, except zero entries for itself.

²¹ We have estimated altogether 26 coefficients (with constant included), 28 industry effects, 17 country effects and 30 year effects.

Also in terms of inter-industry effects, except for industry 6, the results are very similar under two methods. Only industries 3, 4 and 8 have positively significant elasticities in both specifications. Industry 6 has positively significant inter-industry spillovers in OPW but not in OLS even at a 10% level of significance. However, since endogeneity might be an issue behind the OLS estimation, we rely more on OPW results. The most-R&D-intensive industry, industry 1, and the least-R&D-intensive industry, industry 10, have negative inter-industry spillovers in both specifications. Except for few industries, our estimates are not very different from what is in the literature. Keller (2002a) estimates the elasticity of TFP with respect to own-industry R&D between 7% and 17%, whereas the elasticity of domestic R&D from other industries is estimated to be in the range of 5-9%.

For a robustness check, we estimate Equation (9) using labor productivity as a dependent variable with capital-labor ratio as an additional control but without controlling for employment on the right hand side.²² The results are given in Table A11. The results under OPW are very similar under two cases (using value added and labor productivity as dependent variables), except that own R&D coefficient of industry 10 turns negative and insignificant in the latter case. There is no difference in inter-industry effects between two specifications. In the case of OLS estimation, there are two differences. One, like in the OPW case while using labor productivity as dependent variable, the own-effect of industry 10 turns negative and insignificant and second, the inter-industry effect of industry 6 which was significant only at 10% in the former case is significant at 1% level in the latter.

The elasticity on foreign R&D is estimated very consistently in the study. Under OPW, the elasticity of foreign R&D ranges between 0.24-0.29 with intra-industry foreign R&D returns

²² Note that if we control for employment on the right hand side, the estimation results will be the same as those while using value added as a dependent variable, thereby making the estimation redundant.

ranging from 75% to 90%. However, in three of the four cases, foreign inter-industry R&D effect is negatively significant.²³ But since it is insignificant in our preferred estimation (OPW), we conclude that there is no significant inter-industry spillover from foreign R&D. Given that foreign intra-industry R&D effect is highly positive and robust, while estimating return to domestic R&D, one should control for foreign R&D. Otherwise the estimated return to domestic R&D might be biased. However, none of the studies that estimate social returns to R&D have done that.

Since the estimated elasticities are very similar under two very different estimating methods (OPW and OLS) and also similar whether we estimate production function or take labor productivity as a dependent variable, our results are robust. Since the estimation with value added as a dependent variable measures the impact of R&D in TFP, our preferred estimations are given in Table A10. Again, since OPW tackle the potential endogeneity issue in production function estimation, our preferred estimation method is OPW. Hence in the following section, we will use the results under OPW in Table A10 to calculate the rates of returns to R&D, whereas all four set of elasticities in Tables A10 and A11 will be used to develop the ranges of return.

V. Rates of Return to R&D Capital

The results in this paper are derived under the functional form of constant elasticity across country-industry. We can convert these elasticities into rates of return to R&D. The own-industry

²³ There is an asymmetry in the impact of domestic and foreign R&D: in contrast to domestic R&D there are no positive inter-industry effects in foreign R&D. Is this difference just a by-product of our asymmetric treatment of domestic and foreign R&D (in domestic R&D, we allow individual industry effect whereas in foreign we sum all other industries' R&D into one aggregate number using import share as weights) or is it a manifestation of genuine asymmetric impact of domestic and foreign R&D spillovers? If the answer is the latter, then the question is why? Is it because for technology diffusion distance and/or borders are more detrimental in inter-industry than in intra-industry? We do not deal with this issue in this paper (see Keller, 2002b for the importance of distance on technology diffusion).

rate of return to R&D of industry i can be computed as $\rho_i^{or} = \gamma_i \frac{Y_i}{R_i} \equiv \frac{\partial Y_{cif}}{\partial r_{cif}} \frac{Y_{cif}}{R_{cif}}$, and inter-industry

spillovers that industry i supplies to industry j is given by $\rho_i^j = \gamma_i \frac{Y_j}{R_j} \equiv \frac{\partial Y_{cif}}{\partial r_{cif}} \frac{Y_{cif}}{R_{cif}}$, where lower

case letter is the natural log of upper case one. For the purpose of computing rates of return, we will report those corresponding to elasticities that are statistically significant at least at 5% level.

By doing so the impact of an industry whose coefficient is significant economically but insignificant statistically will be ignored. However, there are not many cases like that, the coefficients which are statistically insignificant are also smaller in magnitudes.

The rates of return based on OPW in Tables A10 (columns 1 and 2) are reported in the first and the second columns respectively in Table 2. Column 3 is the sum of columns 2 and 3. As discussed in Griliches and Lichtenberg (1984), the entry in column 3 can be termed as "social net excess rate of return" (SNERR) to R&D capital. The ranges of rates of return based on four specifications (two in Table A10 and two in Table A11—two based on OPW and two based on OLS specifications) are given in the last two columns.

Results show that the own-industry R&D rate of return (part of private return in excess of normal return that R&D labor and capital receives as factors of production) is positive for all industry and varies substantially across industries. The industry with the highest rates of own-industry return is chemical (excluding pharmaceutical) at (57%).²⁴ Besides, office accounting and computing machinery (49%) machinery and equipment (39%), electrical machinery and apparatus (34%) are the three other industries with second, third and fourth largest own-industry rates of return. The next sets of industries in terms of higher own-industry returns are radio, TV and communication (32%), railroad equipment (27%) and pharmaceutical (26%). Motor vehicle

has own-industry R&D return of 20% and medical precision and optical instrument industry has 19%. The lowest own-industry return of 3% is reported for the most-R&D-intensive aircraft and spacecraft industry.

Table 2. Private and social rates of return to R&D capital by industry

	Based on OPW estimation with value added as dependent variable			Range based on four types of estimations
	Own-industry effect	Inter-industry effect	Social net excess rate of returns	Social net excess rate of returns
1. Aircraft and spacecraft	0.03	- 0.15	-0.11	- 0.11 - 0.18
2. Office, accounting & computing machinery	0.49	0	0.49	0.49 - 0.62
3. Radio, TV and communication equip	0.32	0.02	0.34	0.34 - 0.60
4. Pharmaceuticals	0.26	0.69	0.95	0.90 - 1.01
5. Medical, precision and optical instruments	0.19	0	0.19	0.12 - 0.24
6. Motor vehicles, trailers and semi-trailers	0.20	0.04	0.24	0.23 - 0.24
7. Electrical machinery and apparatus, nec	0.34	0	0.34	0.34 - 0.48
8. Chemicals (excluding pharmaceuticals)	0.57	0.33	0.90	0.64 - 0.90
9. Railroad equip. and transport equip n.e.c.	0.27	0	0.27	0.27 - 0.33
10. Machinery and equipment, n.e.c.	0.39	- 0.24	0.16	- 0.58 - 0.16

The rates of return are calculated by multiplying the coefficients in Table A10 by the ratio of industry value added to industry R&D capital stock. As given in Table A6, the mean R&D capital stock for industry 1 is 11627; for industry 2, it is 4474; for industry 3, 13970; for industry 4, 4466, and for industry 5 it is 3374. The R&D stocks for industries 6 through 10 are 8224, 3064, 5654, 252 and 3683. The mean value added for industry 1 is 4137; for industry it is 6770; for industry 3 it is 23849; for industry 4 it is 4776, and for industry 5 it is 6031. For the remaining industries the values are 12364, 7105, 12458, 788, and 16591. To calculate the inter industry effect "other industry" output is computed as average of all industries output 31323 minus the average of each its own industry value added.

This positive own-industry returns means that the rate of return to research labor and research capital is higher than that to ordinary labor and physical capital. Looking at the results, however, there does not seem to have any relationship between R&D intensity and private rates of return to R&D; there is no pattern between these two variables.

Only four industries have inter-industry effect. Pharmaceutical; chemical (excluding pharmaceutical); motor vehicle, and radio, TV and communication industries affect the productivity of other industries positively. The inter-industry effect of the pharmaceutical

²⁴ Note that the ranking of industries based on elasticities in Table A10 may not correspond to the ranking of industries based on rates of return, as value added to R&D capital ratio varies across industries.

industry ranges from 61% to 76% in four specifications and is 69% in the preferred specification. For chemical (excluding pharmaceutical) industry, the spillovers that its R&D send to other industries range from 14% to 33%, with 33% in the preferred estimation. For the motor vehicle industry its inter-industry spillover ranges from 3-7%, whereas that for the radio, TV and communication industry is 2% in all specifications. The aircraft and spacecraft industry and machinery and equipment industries enter with negative inter-industry spillover, whereas for the remaining four industries there are no statistically significant inter-industry spillovers.²⁵

The range of returns we have estimated under four specifications does not vary much; the band is narrow. Nevertheless, the negative inter-industry effects of two industries are somewhat puzzling. In the aerospace industry, however, one needs to consider two facts. First most of the innovations in this sector are product innovations, which, as the research shows, provide lower benefit than the process innovation, to itself and to others. Second, most of the R&D in this sector is for defense purposes; given the tight security and regulations that characterize defense research, it is not easy to imitate (limited technology diffusion) in this industry. Probably, the result would have been different if we had R&D data only on the civic part of aircraft and spacecraft industry. For the machinery and equipment industry, the inter-industry effect is negative and very large. This is the only industry, where preferred point estimates (at least for own-industry effect) are very different from the other estimations. For this industry, we prefer to use point estimates and ignore range estimates. Even then we cannot justify why the inter-industry point estimate is negative and so large.

²⁵ Note that in this set up, having zero inter-industry spillovers from a particular industry does not mean that this industry does not provide inter-industry spillovers to any single industry. What it means is that the average impact of this industry's R&D on remaining 27 industries, on average, is zero. This could happen despite the fact that this particular industry provides spillovers to some industries.

We have estimated part of private return (own-industry effect) and spillover of the 10 most-R&D-intensive industries. We have termed the sum of two as SNERR. If we could calculate the other part of private return that R&D input are paid as normal labor and capital inputs, and add it with SNERR, we would have social returns to R&D. If we assume that labor and capital are perfectly mobile across industries, then similar inputs will be paid the same remuneration. This will allow us to assume that the normal private return to R&D capital is the same (as is implied by our estimation of a single common elasticity for labor and capital across all countries, industries and years). Let us suppose that this normal return is λ across industries. In that case, the social returns to R&D will be λ plus SNERR, and the ranking of industries by own-industry effect will be same as by private return and the ranking of industries by SNERR will be the same as that by social returns to R&D.

Note that with the presence of inter-industry R&D effects, the ranking of industries based on social return to R&D differs from that based on private return to R&D. Hence, the industry that generates the most R&D dollar return for private firms may not be the one that generates the most for society. For example, the pharmaceutical industry, which is seventh-most-rewarding industry in terms of own-industry effect, ranks first in social return. The most-rewarding chemical (excluding pharmaceutical) industry in terms of private return becomes the second rewarding in terms of social return. Despite its nil inter-industry effect, the second most rewarding industry in private return (office accounting and computing) becomes the third-most-socially rewarding industry followed by the radio, TV and communication equipment industry and the electrical machinery and apparatus industry. If we take the negative inter-industry effect seriously, the only industry with negative SNERR is aircraft and spacecraft. However, if the value of λ is greater than 11%, its social return will be positive.

VI. Technology Flow Matrix: R&D Spillovers by Sources and Destinations

We have decomposed the impact of an industry's R&D into itself and to the overall economy which is comprised of the other 27 industries. This task, however, stops short of allowing us to compute the impact of each industry's R&D into the productivity of each other industry in the economy. The focus of this section is to address this issue by developing a technology diffusion flow (or industry-by-industry impact) matrix among source and destination industries of R&D spillovers. To our knowledge, this has not yet been done using production function approach. Using a cost function approach, researchers have estimated inter-industry technology flow matrixes but the number of industries used are a bit aggregate (see Bernstein and Nadiri, 1988). Here we develop a methodology where any number of industries (both as senders and recipients of R&D spillovers) can be used simultaneously with a substantially less number of coefficients to estimate than under a cost function approach. The impact of industry i 's R&D on 28 industries is estimated using the following equation:

$$(10) \quad y_{i,j} = \beta_0 + \beta_1 k_{i,j} + \beta_2 l_{i,j} + \gamma r_{i,i} + \sum_{j=1}^{28} \gamma_j r_{i,j} + \sum_{j=1}^{28} \gamma_j D_j \times r_{i,j} + \phi^1 \sum_{j=1}^{28} m_{i,j} r_{i,j} + \epsilon_{i,j}.$$

Equation (10) is similar to Equation (7) except that the term $\sum_{j=1}^{28} \gamma_j r_{i,j}$ in the latter is decomposed into two terms $\sum_{j=1}^{28} \gamma_j r_{i,j}$ and $\sum_{j=1}^{28} \gamma_j D_j \times r_{i,j}$ in the former. Here $r_{i,i}$ is the R&D of industry i (the industry whose impact on all 28 industries we are estimating); D_j are 28 industry dummies. Our interest is on 28 γ_j coefficients which measure the impact of R&D of industry j on 28 industries (to itself and to others). Again, γ measures the own-industry impact of 18 industries; γ_j measures the total impact of the other 9 industries (among the ten industries except industry i). By estimating equation (10), we try to answer the following question: What is the

impact of industry 1's R&D on each of the 28 industries once the R&D of the other 9 industries as separate variables and all the other 18 industries' R&D are controlled for?

To answer this question, we use an OPW method and report the impact of industries 2, 4, 6 and 8 in Table A12 (industry 1 will be one of these four industries). We take only these four industries because the last three are among the four industries that have positively significant inter-industry effects (Table A10), and industry 2 includes the information and communication technology industry that is regarded to have impacts on many other industries. The column industries are "source industries" and the row industries are "receiving industries" of R&D spillovers. Each column shows how much the column-heading industry spills to each of the row-heading industry (destination of R&D spillover emanated from the column-heading industry). Each row shows how much a row-heading industry receives from each of column-heading industry (sources of R&D spillover gains obtained by row-heading industry). Take column 1: each entry in the column shows the impact of industry 4's R&D on the productivity of row industries. For example, the output elasticity of industry 1 (aircraft and spacecraft) with respect to R&D of industry 4 is 0.283 (entry in first column and first row). Similarly, reading the second column, the output elasticity of industry 1's output with respect to R&D of industry 8 (chemical excluding pharmaceutical) is 0.182. Reading across the row, the impact on industry 10's output of industry 4's R&D is given by 0.164, that of industry 8's R&D by 0.114, and that of industry 6's R&D by 0.051.

Industry 4's R&D has positive impacts on the productivity of all industries except for four services and construction industries (23 through 26 and 28). Industry 8, affects 20 of the 27 industries positively. The R&D of industries 6 affects the productivity of 15 industries positively, and affects industry 2 negatively. The R&D of industry 2 affects 10 industries positively, six

industries negatively and has no effect on the remaining 11 industries. The entries in bold face are own-industry elasticities.

This result provides an answer to the conjecture that Griliches made in 1992 while critically commenting on the methodology of constructing an inter-industry (spillover) effect as a fraction of own-industry effect that was prevalent in the literature. He said, "it is quite possible for an idea to have its entire effect elsewhere than where it was originated (1992, pp. S40-S41)." Based on the results, our answer is that it is not the case that the entire effect of an idea will be somewhere else (in other industries) but the effect of an idea can be higher somewhere else than where it is generated, as inter-industry spillover return to an industry's R&D is found higher than own-industry return.

Using the elasticities in Table A12, we compute the rates of returns to R&D of four industries that are spilled to 10 industries. We could have reported the rates of return spilled to each of 28 industries, however, for brevity, we report only for the 10 most-R&D-intensive industries. These are the group of industries that receive the highest returns from the four source industries. This result perhaps indicates that in order to benefit from other industries' R&D, the recipient industry should have absorptive capacity measured by higher R&D intensity.

Table 3. Technology flow matrix

Receiving industry	Source industry			
	Pharmaceutical (Industry 4)	Chemical (industry 8)	Motor vehicle (Industry 6)	Office accounting and computing (Industry 2)
1. Aircraft and spacecraft	0.24	0.13	0.06	0.05
2. Office, accounting & computing machinery	0.17	0	-0.03	0.39
3. Radio, TV and communication equip	0.43	0	0	0.35
4. Pharmaceuticals	0.15	0.08	0.01	0.02
5. Medical, precision and optical instruments	0.21	0.10	0.03	0
6. Motor vehicles, trailers and semi-trailers	0.50	0.23	0.10	0
7. Electrical machinery and apparatus, nec	0.21	0.13	0.02	0
8. Chemicals (excluding pharmaceuticals)	0.58	0.34	0.09	0.12
9. Railroad equip. and transport equip n.e.c.	0.01	0	0	-0.01
10. Machinery and equipment, n.e.c.	0.57	0.31	0.10	0

Looking at the first column, we see that the rates of return to R&D of pharmaceutical industries to itself reported at 15% (entry in bold face) is lower than that to chemical excluding pharmaceutical (58%), machinery and equipment (57%), motor vehicle (50%) and a few other industries. So, it is possible that the idea generated in one industry might have a larger impact to somewhere else than where it is generated. But not all the ideas are transferred to other industries, as own-industry rate of return is positive to all industries. And also, it is not the case that ideas generated in *all industries* have a larger effect somewhere else, as own-industry return to R&D of industry 8, 6 and 3, is larger than their inter-industry effects.²⁶

VII. Conclusions

Is there additional private return to R&D investment above what the labor and capital are paid as a factor of production in normal price (own-industry effect) making total private return to R&D higher than that of regular labor and capital inputs? Does the R&D of an industry spill benefit to

²⁶ Own rates of return to R&D for all four industries in this table are estimated lower than those in Table 2. The reason for these lower rates could be due to the fact that in the previous estimation we use own industry R&D as endogenous and was represented by its own period lag value in the regression but for Table A12, own industry R&D

other domestic industries—the inter-industry effect? Do the inter-industry spillovers depend on whether industries have upstream-downstream linkages (embodied spillovers) or are they simply disembodied, like a pure public good? Do these own-industry and inter-industry effects vary across industries, and what are their magnitudes? How do the most-R&D-intensive and the largest-R&D-spender industries in OECD countries rank in terms of private and social rates of return to R&D? Are ideas the most beneficial at the industry where they are generated or somewhere else? What is the role of foreign intra-industry and inter-industry R&D effects?

These and several other issues are dealt in this paper using the most comprehensive newly constructed dataset for 28 manufacturing and services industries for 17 OECD countries over the period of 30 years (from 1973 to 2002). Based on the results the answers are that R&D input generates higher private return than other capital inputs in all industries. The rates of private return vary substantially across industries. Only two industries' R&D (pharmaceuticals and motor vehicle) has inter-industry effects.

Neither private return nor inter-industry returns seem to depend on R&D intensity of the industry and R&D size of the industry. The most-R&D-intensive industry (aircraft and spacecraft) has 3% (the lowest) own-industry return, whereas the eighth-largest-R&D-intensive chemical (excluding pharmaceutical) industry has the highest own-industry return of 57%. In terms of size, these industries occupy third and fifth largest shares of global R&D (Table 1). The radio TV and communication equipment industry, which constitutes the largest industry share in global R&D, has 33% own-industry effect. The four industries that have positive inter-industry effects are neither the most R&D intensive nor the biggest R&D spenders.

was treated as exogenous, as there was no own-industry R&D entering separately in the equation and was denoted by industry interaction dummy.

Further extension of the model shows that there is no need to have an arm's-length relationship between industries to have benefits from the R&D of other industries. In other words, R&D benefits seem to be more of a public good nature that can be captured by any industries whether they are related or not in the production process. Therefore, the assumption hitherto made in the literature that R&D inter-industry spillovers are somewhat proportional to the input-output linkage needs to be tempered.

With the presence of the inter-industry spillover, there is a wedge between private and social returns to R&D. As a result of different rates of spillovers across industries, industry ranking based on private return differs from that based on social return to R&D. In some industries, the social rates of return is four-times the private return, whereas for some industries social return is equal to private return as there are no inter-industry spillovers. Hence the magnitude of sub-optimal investment could vary substantially across industries.

The paper provides a comprehensive treatment of sources and destinations heterogeneities of R&D benefits in estimating its social rates of return. Interestingly, sometimes, the industries, which did not invest on R&D, may benefit more than the industry that did it. Therefore, it is hard to trace the path of technology diffusion looking at the production boundary. Results also show that there are asymmetric flows of technology diffusion among industries. There are cases that an industry may absorb a lot of inter-industry R&D spillovers from others but may not be able to generate benefits even from its own R&D, let alone to spill to others.

Finally in the case of foreign R&D, industries benefit only from their own counterpart not from other industries. The inter-industry R&D spillovers that are present at home evaporate at borders.

The results in the paper are robust across different estimation techniques. In the process, it also raises the following important research questions. (1) Why do rates of returns to R&D (both private and social) vary substantially across industries? (2) What determines the magnitude and the asymmetric flows of technology diffusion among industries? (3) Why does distance or do borders seem to matter more for inter-industry than for intra-industry R&D effect? If we could answer these questions, then we will be a step closer in understanding the mechanism of innovation and technology diffusion.

References

- Akerberg, Daniel A; Caves, Kevin and Frazer, Garth.** "Structural Identification of Production Functions", UCLA Working paper, November 2005.
- Arellano, Manuel and Bond, Stephen.** "Some Tests of Specification for Panel Data: Monte Carlo Evidence and an Application to Employment Equations." *Review of Economic Studies*, 1991, 58, pp. 277-97.
- Bernstein, Jeffrey I. and Nadiri, M. Ishaq.** "Interindustry R&D Spillovers, Rates of Return, and Production in High-Tech Industries." *American Economic Review Papers and Proceedings*, May 1988, 78(2), pp. 429-34.
- Blundell, Richard and Bond, Stephen.** "GMM Estimation with Persistent Panel Data: An Application to Production Functions", *Econometric Reviews*, 2000, 19, pp. 321-340.
- Bresnahan, Timothy F.** "Measuring the Spillovers from Technical Advance: Mainframe Computers in Financial Services." *American Economic Review*, September 1986, 76(4), pp. 742-755.
- Clark, Kim B. and Griliches Zvi.** "Productivity Growth and R&D at the Business Level: Results from the PIMS Data Base" in Zvi Griliches, ed., *R&D, Patents and Productivity*. Chicago: UCP, (1984). pp. 393-416.
- Disney, Richard; Haskel, Jonathan and Heden, Ylva.** "Restructuring and Productivity Growth in UK Manufacturing." *Economic Journal*, July 2003, 113, pp. 666-94.
- Coe, David. and Helpman, Elhanan.** "International R&D Spillovers." *European Economic Review*, 1995, 39(5), pp. 859-87.

- Griliches, Zvi.** "Search for R&D Spillovers." *Scandinavian Journal of Economics*, 1992, 94 Supplement pp. S29-47.
- _____. "Issues in Assessing the Contribution of Research and Development to Productivity Growth." *Bell Journal of Economics*, Spring 1979, 10(1), pp. 92-116.
- Griliches, Zvi and Lichtenberg, Frank.** "R&D and Productivity Growth at the Industry Level: Is There Still a Relationship?" in Zvi Griliches, ed., *R&D, Patents and Productivity*. Chicago: University of Chicago Press, 1984, pp. 465-496.
- Griliches, Zvi and Mairesse, Jacques.** "Production Functions: The Search for Identification." *Practicing Econometrics*, edited by Zvi Griliches. Elgar, 1998, pp. 383-411.
- Jaffe, Adam B.** "Technological Opportunity and Spillovers of R&D: Evidence from Firms' Patents, Profits, and Market Value." *American Economic Review*, December 1986, 76(5), pp. 984-1001.
- Jones, Charles I. and Williams, John C.** "Measuring the Social Return to R&D." *Quarterly Journal of Economics*, November 1998, pp. 1119-1135.
- Keller, Wolfgang.** "International Technology Diffusion." *Journal of Economic Literature*, September 2004, XLII, pp. 752-782.
- _____. "Trade and the Transmission of Technology." *Journal of Economic Growth*, March 2002a, 7(1), pp. 5-24.
- _____. "Geographic Localization of International Technology Diffusion." *American Economic Review*, March 2002b, 92(1), pp. 120-42.
- Nadiri, M. Ishaq.** "Innovations and Technological Spillovers." NBER Working Paper No. 4423, August 1993.
- Nickell, Stephen J.** "Competition and Corporate Performance." *Journal of Political Economy*, 1996, 104(4), pp. 724-746.
- Olley, G. Steven and Pakes, Ariel.** "The dynamics of productivity in the telecommunications equipment industry", *Econometrica*, 1996, 64, pp. 1263-1297.
- Schankerman, Mark.** "The Effects of Double-counting and Expensing on the Measured Returns to R&D." *Review of Economics and Statistics*, 1981, 63(3), pp. 454-58.
- Scherer, F. M.** "Inter-Industry Technology Flows and Productivity Growth." *Review of Economics and Statistics*, 1982, 64, pp. 627-634.
- Wooldridge, Jeffrey M.** "On Estimating Firm-level Production Functions Using Proxy Variables to Control for Unobservables", MSU Working paper, April 2004.

Appendix A Table A1: Data availability (Number of Observation)

Appendix A Table A1: Data availability (Number of Observations)						
Country	Name	Value added	Labor	Capital	R&D	
1	Australia (AUS)	840	840	780	785	
2	Belgium (BEL)	840	840	839	420	
3	Canada (CAN)	840	840	840	840	
4	Denmark (DNK)	840	834	810	812	
5	Finland (FIN)	840	840	839	839	
6	France (FRA)	840	837	811	783	
7	Great Britain (GBR)	840	840	837	817	
8	Germany (GER)	840	838	836	807	
9	Ireland (IRL)	840	720	316	661	
10	Italy (ITA)	840	840	840	836	
11	Japan (JPN)	840	840	725	767	
12	S. Korea (KOR)	840	840	840	184	
13	Netherlands (NLD)	838	838	838	786	
14	Norway (NOR)	834	834	834	830	
15	Spain (SPN)	834	724	768	824	
16	Sweden (SWE)	834	834	834	712	
17	USA	833	833	833	706	
International Standard Industrial Classification -3						
					Industry Description	
1	353	480	469	469	465	1. Aircraft and spacecraft
2	30	504	499	499	465	2. Office, accounting and computing machinery
3	32	504	499	499	465	3. Radio, TV and communication equip
4	2423	504	494	494	465	4. Pharmaceuticals
5	33	504	499	499	465	5. Medical, precision and optical instruments
6	34	504	499	499	465	6. Motor vehicles, trailers and semi-trailers
7	31	504	499	499	465	7. Electrical machinery and apparatus, nec
8	24ex2423	504	499	499	465	8. Chemicals (excluding pharmaceuticals)
9	352+359	504	499	499	465	9. Railroad equip. and transport equip n.e.c.
10	29	504	499	499	465	10. Machinery and equipment, n.e.c.
11	23	504	499	499	465	11. Coke, refined petroleum and nuclear fuel
12	25	504	499	499	465	12. Rubber and plastics products
13	272+2732	504	499	499	465	13. Non-ferrous metals
14	271+2731	504	499	499	465	14. Iron and steel
15	351	504	499	499	465	15. Building and repairing of ships and boats
16	26	504	505	505	465	16. Other non-metallic mineral products
17	28	504	499	499	465	17. Fabricated metal products
18	15-16	504	505	505	465	18. Food products, beverages and tobacco
19	36-37	504	505	505	465	19. Manufacturing nec, recycling
20	24-22	504	505	505	465	20. Pulp, paper, printing and publishing
21	17-19	504	505	505	465	21. Textiles, textile products, leather and footwear
22	20	504	505	505	465	22. Wood and products of wood and cork
23	72-74	505	505	505	369	23. Computer, R&D and other business
24	64	506	506	506	300	24. Post and telecommunications
25	60-63	506	506	506	332	25. Transport and storage
26	50-52, 55, 65-67, 70, 75-99	504	504	504	362	26. Other services
27	40-41	505	506	506	411	27. Electricity, gas and water supply
28	45	506	506	506	405	28. Construction

Table A2: R&D Intensity (R&D expenditure over industry value added in %), 1973-2002

Country	AUS	BEL	CAN	DNK	FIN	FRA	GBR	GER	IRL	ITA	JPN	KOR	NLD	NOR	SPN	SWE	USA	Industry share
Industry																		
1	1.6	18.4	19.2	0.0	2.3	57.2	26.6	50.3		26.5	23.9	87.5	13.0	2.7	28.1	32.5	40.1	38.3
2	9.8	5.9	38.2	16.8	15.4	14.9	11.3	14.0	2.0	23.8	29.6	37.3	97.3	31.9	6.9	16.4	40.3	30.5
3	20.3	51.2	43.6	17.8	29.1	33.1	20.1	40.0	16.0	18.0	15.9	57.5	10.8	39.9	11.1	52.3	28.9	26.8
4	4.8	40.5	17.0	30.3	30.4	24.4	40.2	20.8	5.0	9.4	18.6	9.4	24.4	8.4	8.5	40.4	22.6	21.2
5	15.1	18.4	3.8	14.7	14.6	18.1	6.9	6.6	3.0	2.7	16.8	16.4	7.2	15.6	4.0	18.9	15.5	14.4
6	6.4	3.5	1.1	4.0	3.4	12.2	8.5	12.4	5.5	9.6	26.5	36.7	7.9	7.0	2.7	19.3	15.0	13.3
7	3.4	8.2	3.9	5.3	10.2	5.8	9.1	6.3	2.9	2.7	16.6	16.9	45.5	6.9	2.2	9.9	7.4	8.9
8	4.7	13.9	2.0	6.0	7.4	7.9	6.6	11.4	0.5	3.3	13.8	12.8	9.2	10.0	2.1	5.8	8.1	8.9
9	3.7	19.0	1.3	8.8	10.4	5.1	7.9	8.6	0.4	3.4	6.9	12.3	2.1	1.9	3.7	7.0	12.6	8.1
10	3.5	7.1	2.0	5.8	6.3	3.8	4.3	5.1	2.6	1.5	6.9	13.8	4.7	6.9	2.0	9.3	4.1	4.8
11	1.0	4.4	8.5	1.0	4.5	3.7	9.4	1.9	0.0	1.3	2.7	4.4	5.1	3.7	1.0	2.4	7.3	5.0
12	1.1	5.4	0.8	2.0	4.2	4.1	0.9	2.4	2.3	1.4	5.6	6.3	1.8	2.7	1.3	3.3	3.2	3.2
13	2.2	5.4	3.5	0.0	7.4	4.1	1.7	2.0	1.1	0.9	5.7	4.1	5.2	5.0	0.7	3.1	2.1	3.2
14	4.1	2.9	0.6	2.9	2.1	3.1	1.5	1.7	1.8	0.9	3.7	2.7	8.8	4.4	0.9	4.1	1.4	2.4
15	5.2	1.4	0.0	4.3	2.3	1.2	2.7	2.3	2.2	2.2	1.5	5.8	1.0	2.1	4.0	3.1	2.5	2.8
16	1.3	3.0	0.4	1.3	2.3	2.2	1.1	1.6	1.6	0.2	4.5	3.3	0.6	1.6	0.5	1.8	2.3	2.2
17	0.8	2.8	1.0	0.9	2.5	0.8	0.8	1.5	1.9	0.4	1.9	2.9	0.9	1.6	0.5	1.9	1.5	1.3
18	1.2	1.6	0.5	1.6	2.0	0.9	1.2	0.6	0.9	0.3	2.3	2.9	2.0	1.6	0.4	1.7	1.4	1.3
19	0.8	2.5	1.0	4.3	1.3	0.9	0.8	0.6	0.6	0.1	1.7	4.2	0.5	1.3	0.5	0.8	1.4	1.2
20	0.9	1.6	0.9	0.2	1.4	0.3	0.5	0.3	0.2	0.1	0.9	1.6	0.2	1.3	0.3	2.0	1.2	0.9
21	0.4	2.3	0.9	0.4	1.0	0.7	0.4	0.9	1.6	0.1	1.4	1.5	0.7	1.3	0.3	1.2	0.5	0.7
22	0.5	1.0	0.4	0.4	1.0	0.3	0.1	0.8	1.7	0.1	1.8	0.8	0.3	1.0	0.1	0.4	0.5	0.6
23	0.1	0.1	0.1	0.7	1.6	0.1	0.1	0.1	5.9	0.7	0.4	3.9	2.4	17.9	4.2	25.6	0.6	0.9
24	0.0	0.1	0.1	0.2	0.5	0.1	0.1	0.2	2.2	0.7	0.4	2.0	0.5	16.1	3.7	7.0	1.2	1.0
25	0.3	0.3	0.2	0.8	0.5	0.1	0.1	0.1	1.8	0.1	0.0	0.3	0.3	1.0	0.5	1.0	0.2	0.2
26	0.0	0.0	0.0	0.0	0.1	0.0	0.0	0.0	0.5	0.0	0.0	0.1	0.1	1.6	0.5	3.5	0.0	0.1
27	0.1	0.2	0.3	0.4	4.1	0.4	0.3	0.4	4.0	1.0	0.2	2.1	0.3	5.5	1.2	1.9	0.1	0.4
28	0.2	0.3	0.2	0.9	1.2	0.1	0.2	0.1	3.0	0.3	0.0	0.6	0.9	5.5	0.4	2.9	1.6	0.7
Avg.	0.5	1.7	0.7	0.9	1.9	1.4	1.2	1.7	2.2	0.7	2.0	5.0	1.1	4.0	1.0	6.2	1.8	1.7

Note: The industry names corresponding to numbers reported in column 1 are as given in Table A1

Table A3: Industry Share of R&D Expenditure by Country and at the Aggregate (in %), 1973-2002

Country	AUS	BEL	CAN	DNK	FIN	FRA	UK	GER	IRL	ITA	JPN	KOR	NLD	NOR	SPN	SWE	USA	Industry share
Industry																		
1	0.5	1.7	13.7	0.0	0.1	16.2	14.0	7.1	0.1	9.5	0.8	2.0	1.1	0.1	3.1	1.6	16.3	9.8
2	1.8	0.3	6.3	1.6	1.4	3.4	3.7	2.9	1.7	3.8	8.9	4.1	8.9	0.7	1.2	0.6	9.4	6.8
3	12.8	20.1	34.3	6.8	32.0	14.4	12.2	14.3	9.8	15.5	17.2	40.9	8.2	3.2	4.2	8.7	13.9	16.0
4	6.3	18.3	7.2	21.8	4.6	10.9	20.0	6.0	3.8	9.4	6.6	1.8	6.0	1.2	4.3	5.5	6.9	7.2
5	5.4	1.8	2.0	8.1	3.2	9.1	3.6	3.5	1.8	2.2	3.8	1.1	1.7	1.0	0.9	1.7	9.1	5.8
6	13.1	3.1	3.2	1.0	0.6	13.3	8.8	21.4	0.4	13.5	13.1	18.0	2.2	0.3	4.3	6.4	10.9	11.9
7	2.5	3.7	2.3	3.2	5.3	3.8	6.1	7.5	1.2	4.3	10.5	2.1	9.7	1.0	1.8	1.0	3.2	5.0
8	6.6	22.0	3.8	4.2	5.2	8.3	8.2	14.5	1.4	6.0	10.0	5.6	14.7	1.6	2.5	0.9	6.3	7.9
9	1.0	0.6	0.3	0.3	0.7	0.4	0.6	0.5	0.1	0.8	0.3	0.3	0.1	0.0	0.4	0.2	0.6	0.5
10	6.3	5.8	3.5	13.4	10.0	4.8	7.0	11.2	1.0	6.2	8.9	4.5	4.2	2.1	2.6	4.0	3.8	5.8
11	0.6	1.2	3.3	0.1	1.2	2.0	3.4	0.4	0.0	0.8	0.8	1.4	1.7	0.2	0.6	0.1	1.9	1.4
12	1.6	2.1	0.9	1.5	1.8	2.9	0.7	1.6	0.6	2.0	2.6	1.4	0.7	0.2	1.1	0.3	1.2	1.5
13	3.5	1.3	3.0	0.0	0.9	0.7	0.3	0.4	0.0	0.3	1.6	0.3	0.2	1.1	0.2	0.1	0.4	0.7
14	6.7	2.4	0.6	0.3	1.1	1.3	0.7	1.0	0.1	0.9	2.9	1.3	1.0	0.4	0.6	0.6	0.4	1.1
15	1.8	0.0	0.0	1.5	0.7	0.1	0.4	0.2	0.0	0.5	0.2	1.7	0.2	0.8	0.7	0.1	0.2	0.3
16	2.2	2.0	0.3	1.1	1.1	1.3	0.7	1.0	0.6	0.4	2.3	1.0	0.3	0.2	0.7	0.2	0.7	1.0
17	2.2	2.4	1.7	1.2	2.1	1.1	0.9	2.0	0.5	1.3	1.4	0.8	0.9	0.4	0.8	0.5	1.1	1.2
18	7.6	2.5	1.9	5.2	2.8	1.9	2.7	0.8	2.8	1.1	2.6	1.5	4.7	0.9	1.3	0.5	1.3	1.7
19	0.9	1.0	0.9	4.3	0.5	0.6	0.4	0.3	0.2	0.3	0.9	0.5	0.4	0.2	0.4	0.1	0.5	0.6
20	3.7	1.4	2.9	0.4	4.2	0.3	1.1	0.3	0.3	0.2	0.9	0.5	0.3	0.7	0.4	1.1	1.4	1.0
21	0.8	1.9	1.1	0.3	0.5	0.6	0.4	0.6	0.6	0.4	0.9	0.8	0.3	0.1	0.6	0.1	0.3	0.5
22	0.6	0.2	0.6	0.2	0.7	0.1	0.0	0.2	0.2	0.1	0.3	0.0	0.0	0.2	0.0	0.1	0.2	0.2
23	1.6	0.7	0.8	4.1	4.4	0.6	0.9	0.7	32.4	9.3	1.3	3.6	14.7	28.6	21.3	27.6	2.9	4.2
24	0.2	0.3	0.3	1.9	0.8	0.1	0.2	0.2	8.2	2.2	0.3	0.9	4.7	15.4	9.6	4.6	2.2	1.8
25	3.7	1.0	1.2	5.0	2.1	0.2	0.3	0.1	2.5	0.6	0.1	0.3	1.1	2.3	2.8	1.0	0.4	0.5
26	1.9	0.7	1.2	1.3	1.4	0.3	0.5	0.4	6.9	2.2	0.3	1.1	2.8	21.3	26.8	28.4	0.2	2.2
27	0.9	0.9	1.3	6.0	6.1	0.8	1.2	0.5	13.8	3.4	0.3	1.2	4.8	7.9	3.8	1.8	0.2	0.9
28	3.2	0.9	1.4	5.2	4.3	0.4	0.8	0.4	8.9	2.8	0.2	1.2	4.4	8.0	2.9	2.4	4.3	2.5
Total Country's share	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100
	0.5	1.2	1.5	0.4	0.5	5.2	4.3	9.5	0.5	2.2	18.0	6.1	1.7	1.2	1.9	3.6	41.8	

Note: The industry names corresponding to numbers reported in column 1 are as given in Table A1.

Table A4: Employment by Country & Industry (Total number of workers engaged; in 1000), 1973-2002

Country	AUS	BEL	CAN	DNK	FIN	FRA	UK	GER	IRL	ITA	JPN	KOR	NLD	NOR	SPN	SWE	USA	Average
Industry																		
1	14	6	36	1	3	86	168	69		35	29	5	11	3	11	13	696	74
2	9	1	15	2	3	65	54	91	12	25	249	34	8	2	12	7	305	52
3	26	26	40	12	18	219	207	235	8	119	998	318	84	7	36	37	686	181
4	24	15	17	10	3	72	82	115	4	87	152	49	12	7	42	14	262	57
5	15	7	46	14	6	129	154	371	12	108	301	57	16	5	33	22	945	132
6	75	60	132	8	8	341	326	789	5	260	300	200	27	4	176	69	891	216
7	39	38	56	19	18	305	234	614	12	246	771	139	15	12	87	34	905	209
8	31	57	77	14	17	186	258	526	14	186	343	112	76	9	101	25	795	166
9	18	4	11	2	4	27	34	51	5	38	48	15	5	3	19	6	100	23
10	62	51	117	73	62	416	564	1361	14	560	1410	288	85	28	166	109	1936	429
11	6	7	17	1	4	41	34	44	1	29	47	31	8	2	11	3	165	27
12	46	24	80	20	16	205	243	367	9	174	409	168	30	8	98	26	839	162
13	35	13	50	6	3	48	71	125	1	39	154	21	3	13	23	11	331	56
14	45	52	57	3	15	152	196	377	1	152	443	115	10	6	71	40	544	134
15	14	2	19	14	16	41	93	59	1	41	155	85	29	35	42	14	197	50
16	54	45	56	25	20	195	214	369	13	316	607	160	37	12	184	26	605	173
17	113	69	137	47	33	464	513	863	13	626	1138	199	107	20	246	87	1496	363
18	180	106	260	92	56	617	610	1008	57	474	1513	290	172	55	412	74	1769	456
19	74	41	109	34	22	233	193	350	11	310	1025	174	115	14	185	61	1054	236
20	127	60	253	57	86	346	505	650	21	293	1046	188	149	51	172	118	2209	372
21	114	105	179	34	50	511	608	629	30	1196	1872	1034	58	16	422	29	1975	521
22	48	15	114	14	40	116	126	210	5	229	433	64	27	23	106	48	776	141
23	486	338	917	154	106	1755	2511	2084	61	1169	3141	413	675	115	580	241	9814	1445
24	172	378	638	417	59	656	1214	698	132	316	649	178	1869	140	201	263	2145	596
25	359	205	553	154	118	900	1165	1402	55	797	3027	706	431	145	632	209	3677	855
26	3911	2018	7099	1366	1104	11184	13743	15818	543	9885	28123	6764	3504	1081	6126	2280	72102	10979
27	139	438	616	469	40	504	1207	480	144	193	702	136	2339	145	155	259	873	520
28	535	248	823	197	168	1629	1922	2632	106	1547	5947	1162	547	139	1283	267	6306	1497
Avg.	242	158	447	116	75	766	973	1157	48	695	1965	468	373	75	415	157	4086	20122

Note: The industry names corresponding to numbers reported in column 1 are as given in Table A1

Table A5: Capital Stock by Country & Industry (in billions US\$ PPP 1995; depreciation rate 5%, 1973-2002)

Country	AUS	BEL	CAN	DNK	FIN	FRA	UK	GER	IRL	ITA	JPN	KOR	NLD	NOR	SPN	SWE	USA	Average
Industry																		
1		21.9	2.2	0.0	9.1	13	11	3		2.4	30	0.5	0.4	0.1	0.3	1.0	45	10
2	0.3	0.0	0.4	0.0	0.0	2	2	3	2.6	1.3	19	3.0	0.3	0.0	0.1	0.2	16	3
3	0.4	3.6	3.5	0.5	0.9	6	10	22	4.0	7.5	77	38.6	6.1	0.3	1.2	2.3	110	17
4	27.5	0.9	1.7	1.3	0.5	4	7	13	0.9	67.5	21	1.8	1.4	0.3	4.1	2.1	30	11
5	0.5	0.4	1.1	0.5	0.3	14	5	21	0.5	4.4	14	0.9	1.5	0.5	0.2	0.5	46	37
6	4.6	52.5	14.8	0.2	0.5	51	28	87	0.0	35.9	128	13.4	3.4	0.4	12.2	8.8	151	7
7	3.2	5.3	4.7	0.7	1.5	18	14	27	0.4	14.4	37	7.7	0.7	1.6	2.6	1.2	24	10
8	4.5	11.8	30.1	2.1	3.0	22	28	56	1.7	27.4	132	16.0	16.7	2.7	9.2	3.8	183	32
9		0.1	2.0		0.8	1	1	4		3.3	10	121.2	0.2	0.1	0.4	0.3	3	10
10	14.7	2.7	3.8	4.1	3.1	21	101	100	0.3	59.9	89	7.7	5.4	1.3	7.7	6.1	84	30
11	2.1	4.3	9.1	0.4	2.0	120	17	157		154.2	76	7.1	11.3	1.3	467.5	1.3	82	70
12	4.4	3.9	5.8	1.7	0.9	15	21	23	0.3	26.6	102	9.9	4.7	1.0	69.1	2.0	44	20
13	12.2	1.7	16.2	10.2	0.5	17	4	15		6.1	41	3.0	5.9	2.1	1.1	0.9	43	11
14	13.1	15.1	59.1	0.3	1.9	55	299	28		35.4	153	21.8	5.1	1.7	4.7	5.3	106	50
15	0.7	0.1	0.5	1.6	3.2	0	7	2		6.5	38	6.4	1.4	3.2	0.6	5.3	18	6
16	4.8	5.3	11.8	5.1	3.2	36	97	50	0.3	38.0	123	9.3	6.0	2.4	290.1	4.8	73	45
17	4.4	3.0	10.6	2.9	1.4	119	31	46	0.2	64.3	67	8.9	8.0	1.8	4.3	3.9	82	27
18	16.6	13.0	25.9	7.2	6.0	87	72	91	1.9	51.0	185	17.6	20.1	9.0	281.1	13.2	209	65
19	0.3	21.0	3.3	0.7	1.3	18	36	66	0.3	24.2	13	1.8	3.7	1.1	5.5	0.7	47	14
20	5.4	4.7	56.8	6.7	19.3	42	47	65	0.7	22.5	181	8.0	10.0	6.1	12.0	31.0	302	48
21	3.0	7.5	8.5	1.9	3.7	49	149	55	0.2	81.9	297	30.9	4.5	2.5	48.5	11.0	67	48
22	4.3	1.5	15.7	6.0	3.4	32	6	29	0.1	24.6	36	4.0	2.2	1.9	60.4	8.2	197	25
23	52.7	175.8	50.9	9.2	18.4	508	6442	278	38.3	8476.6	60	32.5	41.5	5.7	31.9	36.9	597	992
24	27.6	14.0	84.8	37.2	22.5	346	104	57	19.4	114.6	1164	1.2	2559.2	20.6	29.9	25.7	590	307
25	73.8	32.2	420.5	37.3	88.1	622	215	121	9.5	238.2	7	1.4	64.2	38.2	157.1	36.5	484	156
26	543.7	233.5	443.8	106.2	126.5	1788	1051	1388	19.2	5155.7	7	267.8	643.8	1222.7	1240.5	92.6	49200	3737
27	53.4	104.9	282.5	106.8	19.6	201	156	314	21.0	360.0	15100	50.3	230.1	47.9	171.1	66.4	840	1072
28	23.9	23.1	28.0	6.4	6.4	182	37	258	2.9	93.3	11000	27.9	34.1	4.3	96.4	11.0	321	715
Average	35	27	57	13	12	157	321	121	6	543	1043	26	132	49	107	14	1932	270

Note: The industry names corresponding to numbers reported in column 1 are as given in Table A1.

Table A6: R&D Capital Stock by Country & Industry (in millions of US \$ PPP 1995, with 15% depreciation rate), 1973-2002

Industry	AUS	BEL	CAN	DNK	FIN	FRA	GBR	GER	IRL	ITA	JPN	KOR	NLD	NOR	SPN	SWE	USA	Average
1	40	164	3155	0	12	17188	13097	9333	4	2476	1454	413	230	6	540	903	148638	11,627
2	109	55	1092	47	51	1338	773	2393	154	475	21709	7246	2943	40	204	174	37251	4,474
3	888	2984	6365	290	1201	10005	7222	17525	1150	4377	37832	58261	2144	497	992	4149	81607	13,970
4	395	1642	958	807	224	4933	7867	6262	182	2754	11124	570	1204	174	727	2056	34051	4,466
5	229	213	253	313	165	5004	1681	2743	68	305	5872	259	237	124	140	459	39287	3,374
6	719	327	454	166	34	9379	4930	20788	23	4824	20511	6722	384	33	1002	2175	67340	8,224
7	359	476	493	106	320	2170	4457	10418	61	1011	13536	660	3032	196	338	473	13989	3,064
8	421	2404	1028	180	304	5138	4667	16772	77	1650	19293	2070	3169	279	484	473	37713	5,654
9	53	77	107	20	37	254	127	364	4	245	514	54	9	12	84	94	2223	252
10	447	636	474	536	532	2629	4528	14884	49	1324	15476	1282	618	285	406	1814	16686	3,683
11	29	125	657	3	46	2393	2930	707	1	2873	2456	339	1957	27	956	25	12509	1,649
12	134	193	140	59	86	1756	501	1659	32	662	5632	425	126	35	649	136	5297	1,031
13	203	142	783	2	67	565	285	521	2	97	2769	123	85	232	37	80	3146	538
14	347	281	185	11	78	851	763	1101	4	350	5941	528	202	95	128	295	2864	825
15	80	5	1	104	56	68	295	203	2	176	473	448	31	180	135	163	5456	463
16	136	194	102	90	74	973	948	1148	28	104	4652	337	60	38	326	131	4111	791
17	149	232	253	59	92	862	774	2122	33	442	2639	235	154	88	138	351	5695	842
18	409	240	476	198	167	1096	2087	954	158	226	5614	396	922	145	565	356	7444	1,262
19	42	101	136	303	20	260	903	230	9	118	1741	135	131	17	54	26	2448	393
20	158	158	765	20	350	285	570	359	15	35	2290	159	62	123	69	692	7285	788
21	48	186	172	16	35	603	784	580	36	47	2166	231	61	16	90	47	1349	380
22	42	20	113	12	44	102	117	335	5	15	654	9	21	33	12	31	2191	221
23	1231	589	2421	540	167	2009	7239	1689	67	2706	272	834	685	358	788	941	35420	3,409
24	547	211	697	208	149	1711	2376	8336	69	341	2333	2019	391	133	341	506	11612	1,881
25	119	124	400	381	34	474	257	806	32	22	894	831	605	48	111	82	2035	427
26	878	220	1278	754	174	2944	333	744	61	328	52	1591	473	151	55	230	70989	4,780
27	147	106	1106	66	104	1509	1069	790	55	1681	2444	892	406	72	264	306	1092	712
28	319	297	511	166	82	1084	746	843	36	564	5552	2165	306	136	236	294	12936	1,545
Avg.	310	443	878	195	168	2771	2583	4450	86	1080	6996	3187	737	128	353	624	24024	2883

Note: The industry names corresponding to numbers reported in column 1 are as given in Table A1.

Table A7. Effects of domestic and foreign R&D at the aggregate level

	(1) OLS	(2) OLS	(3) OLS	(4) OLS	(5) IV System GMM	(6) Olley- Pakes	(7) OPW
Labor	0.645 (0.015) ^a	0.496 (0.015) ^a	0.544 (0.011) ^a	0.358 (0.023) ^a	0.565 (0.037) ^a	0.482 (0.037) ^a	0.491 (0.009) ^a
Capital	0.273 (0.011) ^a	0.268 (0.009) ^a	0.212 (0.007) ^a	0.415 (0.021) ^a	0.207 (0.019) ^a	0.128 (0.117)	0.193 (0.022) ^a
Domestic R&D		0.197 (0.007) ^a	0.185 (0.006) ^a	0.157 (0.014) ^a	0.183 (0.021) ^a	0.156 (0.023) ^a	0.154 (0.004) ^a
Foreign R&D			0.383 (0.011) ^a	0.297 (0.014) ^a	0.343 (0.035) ^a	0.276 (0.038) ^a	0.258 (0.006) ^a
Country FE	Yes	Yes	Yes	No	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes	No	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Country*industry FE	No	No	No	Yes	No	No	No
R square	0.929	0.945	0.953				
OverID p-value					0.720		
AR(1) p-value					0.013		
AR(2) p-value					0.830		
N	13,376	11,980	10,176	10,176	9,143	9,762	9,850
CRS p-value	0.00	0.00	0.00	0.00	0.00	0.00	0.00

Standard errors in parentheses; all are robust standard errors except in column (6) and (7) in which case, robust estimation is not allowed. The instruments are $l_{i,t-2}$, $l_{i,t-3}$ for labor (l) $k_{i,t-2}$ for capital (k) and $r_{i,t-2}$ for domestic R&D (r) in IV System GMM. For OPW specification, labor and domestic R&D are represented by their one period lagged values. In all regressions, foreign R&D is treated as exogenous.

In all regressions, country-, industry-, and time fixed effects are included.

a = significant at the 1% level, b = significant at the 5% level

Table A8. Inter-industry domestic R&D effects (value added as dependent variable)

	(1) OLS	(2) System GMM	(3) OPW	(4) OLS	(5) System GMM	(6) OPW
Labor	0.570 (0.011) ^a	0.593 (0.036) ^a	0.499 (0.009) ^a	0.652 (0.012) ^a	0.694 (0.041) ^a	0.563 (0.009) ^a
Capital	0.196 (0.006) ^a	0.191 (0.017) ^a	0.146 (0.022) ^a	0.270 (0.008) ^a	0.257 (0.025) ^a	0.276 (0.023) ^a
Own industry R&D of all 28 industries	0.168 (0.006) ^a	0.164 (0.018) ^a	0.143 (0.004) ^a			
Own industry R&D of 18 industries				0.013 (0.005) ^b	-0.009 (0.015)	0.051 (0.004) ^a
Inter-industry effects				Total (own plus inter-industry effects)		
Industry 1's R&D	-0.031 (0.008) ^a	-0.030 (0.029)	-0.026 (0.005) ^a	-0.067 (0.010) ^a	-0.064 (0.020) ^a	-0.051 (0.008) ^a
Industry 2's R&D	-0.141 (0.007) ^a	-0.139 (0.019) ^a	-0.071 (0.006) ^a	0.011 (0.010)	0.008 (0.020)	0.010 (0.008)
Industry 3's R&D	-0.016 (0.011)	-0.031 (0.023)	0.024 (0.007) ^a	0.070 (0.012) ^a	0.071 (0.028) ^b	0.049 (0.010) ^a
Industry 4's R&D	0.023 (0.010) ^a	0.023 (0.035)	0.048 (0.009) ^a	0.143 (0.022) ^a	0.111 (0.050) ^b	0.130 (0.018) ^a
Industry 5's R&D	0.037 (0.008) ^a	0.029 (0.027)	0.017 (0.007) ^b	0.012 (0.013)	0.013 (0.028)	0.001 (0.011)
Industry 6's R&D	0.037 (0.006) ^a	0.036 (0.017) ^b	0.029 (0.005) ^a	0.026 (0.011) ^b	0.031 (0.023)	0.029 (0.008) ^a
Industry 7's R&D	0.006 (0.008)	0.009 (0.019)	-0.011 (0.007)	-0.005 (0.015)	-0.015 (0.038)	-0.014 (0.012)
Industry 8's R&D	-0.009 (0.009)	-0.001 (0.026)	-0.008 (0.009)	0.068 (0.030) ^b	0.086 (0.059)	0.091 (0.024) ^a
Industry 9's R&D	0.015 (0.007) ^b	0.015 (0.017)	0.022 (0.006) ^a	0.001 (0.010)	-0.003 (0.022)	0.005 (0.008)
Industry 10's R&D	0.002 (0.008)	0.005 (0.024)	-0.002 (0.010)	-0.119 (0.028) ^a	-0.093 (0.061)	-0.077 (0.022) ^a
Foreign R&D	0.323 (0.011) ^a	0.293 (0.032) ^a	0.254 (0.007) ^a	0.453 (0.013) ^a	0.415 (0.041) ^a	0.289 (0.007) ^a
Over ID p-value		0.800			0.986	
AR(1) p-value		0.030			0.010	
AR(2) p-value		0.787			0.657	
R square	0.958			0.947		
N	10,176	9,143	9,805	10,176	9,244	9,856

For GMM in column (2), the instruments are l_{t-2} , l_{t-3} , k_{t-2} and r_{t-2} for labor, capital and own R&D respectively. The inter-industry R&D of ten industries is taken as exogenous. For OPW in column (3), labor (l) and own R&D (r) are considered endogenous and are represented by their one period lagged values. For GMM in column (5), only l_{t-2} , l_{t-3} , and k_{t-2} are taken as instruments and 18 industries' own R&D and ten industries' inter-industry R&D are considered exogenous. For OPW in column (6), only labor is considered endogenous and is represented by its one period lagged value. In all regressions, country-, industry-, and time fixed effects are included. a = significant at the 1% level, b = significant at the 5% level.

Table A9. Inter-industry R&D effects through input-output linkages

	OPW (1)	OPW (2)	OPW (3)	OPW (4)	OLS (5)
Labor	0.514 ^a	0.518 ^a	0.519 ^a	0.507 ^a	0.579 ^a
Capital	0.257 ^a	0.086 ^a	0.084 ^a	0.104 ^a	0.221 ^a
Own industry R&D of 18 industries	0.059 ^a	0.062 ^a	0.062 ^a	0.085 ^a	0.094 ^a
Industry 1's R&D	-0.022 ^a	-0.028 ^a	-0.030 ^a	-0.055 ^a	-0.055 ^a
Industry 2's R&D		0.003	-0.010	0.000	0.001
Industry 3's R&D			0.057 ^a	0.039 ^a	0.036 ^a
Industry 4's R&D				0.112 ^a	0.118 ^a
Industry 5's R&D				-0.018	-0.006
Industry 6's R&D				0.018 ^b	0.008
Industry 7's R&D				-0.001	0.015
Industry 8's R&D				0.084 ^a	0.029
Industry 9's R&D				-0.002	0.002
Industry 10's R&D				-0.044 ^b	-0.071 ^a
Industry 1's input share × industry 1's R&D	0.139 ^a	0.150 ^b	0.156 ^a	0.200 ^a	0.244 ^a
Industry 2's input share × industry 2's R&D		0.616 ^a	0.622 ^a	0.640 ^a	0.773 ^a
Industry 3's input share × industry 3's R&D			0.062 ^b	0.143 ^a	0.459 ^a
Industry 4's input share × industry 4's R&D				0.497 ^a	0.759 ^a
Industry 5's input share × industry 5's R&D				0.234 ^a	0.186 ^a
Industry 6's input share × industry 6's R&D				0.208 ^a	0.273 ^a
Industry 7's input share × industry 7's R&D				0.170 ^a	0.240 ^a
Industry 8's input share × industry 8's R&D				0.176 ^a	0.243 ^a
Industry 9's input share × industry 9's R&D				0.075	0.068
Industry 10's input share × industry 10's R&D				0.149 ^a	0.217 ^a
Foreign R&D	0.289 ^a	0.245 ^a	0.249 ^a	0.246 ^a	0.311 ^a
N	9,340	9,340	9,340	9,340	9,644

Only labor is considered endogenous and is represented by its one period lagged value. Coefficients on labor and capital are estimated but not reported. Standard errors also are not reported. In all specifications, foreign R&D is treated as exogenous. In all regressions, country-, industry-, and time fixed effects are included. a = significant at the 1% level; b = significant at the 5% level.

Table A10. Own- and inter-industry domestic R&D effects (value added as dependent variable)

	OPW		OLS	
	Own-industry effect	Inter-industry effect	Own-industry effect	Inter-industry effect
Labor	0.507 (0.009) ^a		0.575 (0.012) ^a	
Capital	0.069 (0.023) ^a		0.178 (0.006) ^a	
Own industry R&D of 18 industries	0.133 (0.005) ^a		0.141 (0.006) ^a	
Industry 1's R&D	0.096 (0.010) ^a	- 0.063 (0.008) ^a	0.097 (0.014) ^a	- 0.079 (0.008) ^a
Industry 2's R&D	0.327 (0.013) ^a	0.006 (0.008)	0.407 (0.014) ^a	0.002 (0.008)
Industry 3's R&D	0.188 (0.015) ^a	0.028 (0.010) ^a	0.342 (0.022) ^a	0.039 (0.010) ^a
Industry 4's R&D	0.241 (0.021) ^a	0.116 (0.018) ^a	0.272 (0.022) ^a	0.102 (0.019) ^a
Industry 5's R&D	0.105 (0.015) ^a	- 0.018 (0.011)	0.134 (0.018) ^a	- 0.009 (0.012)
Industry 6's R&D	0.132 (0.011) ^a	0.019 (0.008) ^b	0.155 (0.013) ^a	0.015 (0.009)
Industry 7's R&D	0.146 (0.016) ^a	- 0.019 (0.012)	0.205 (0.016) ^a	- 0.008 (0.013)
Industry 8's R&D	0.258 (0.024) ^a	0.098 (0.022) ^a	0.225 (0.026) ^a	0.044 (0.024) ^b
Industry 9's R&D	0.085 (0.014) ^a	- 0.004 (0.008)	0.104 (0.013) ^a	- 0.002 (0.009)
Industry 10's R&D	0.087 (0.024) ^a	- 0.059 (0.021) ^a	0.083 (0.025) ^a	- 0.094 (0.024) ^a
Foreign R&D	0.238 (0.007) ^a	- 0.031 (0.023)	0.271 (0.012) ^a	- 0.064 (0.026) ^b
R square			0.962	
N	9,805		10,176	

For OPW specification, labor and own-industry domestic R&D of all industries are considered endogenous and are represented by their one period lagged values. The inter-industry R&D is considered as exogenous. Foreign R&D is treated as exogenous. In all regressions, country-, industry-, and time fixed effects are included.

a = significant at the 1% level, b = significant at the 5% level.

Table A11. Own- and inter-industry domestic R&D effects
(labor productivity as dependent variable)

	OPW		OLS	
	Own-industry effect	Inter-industry effect	Own-industry effect	Inter-industry effect
Own industry R&D of 18 industries	0.098 (0.004) ^a		0.095 (0.007) ^a	
Capital-labor ratio	0.198 (0.04) ^a		0.277 (0.013) ^a	
Industry 1's R&D	0.047 (0.011) ^a	- 0.082 (0.008) ^a	0.043 (0.013) ^a	- 0.085 (0.009) ^a
Industry 2's R&D	0.356 (0.011) ^a	0.006 (0.009)	0.355 (0.015) ^a	0.002 (0.009)
Industry 3's R&D	0.234 (0.015) ^a	0.018 (0.011) ^c	0.270 (0.021) ^a	0.036 (0.011) ^a
Industry 4's R&D	0.241 (0.023) ^a	0.128 (0.019) ^a	0.240 (0.023) ^a	0.125 (0.020) ^a
Industry 5's R&D	0.073 (0.016) ^a	0.005 (0.012)	0.069 (0.018) ^a	0.006 (0.013)
Industry 6's R&D	0.113 (0.011) ^a	0.030 (0.009) ^a	0.110 (0.012) ^a	0.029 (0.009) ^a
Industry 7's R&D	0.175 (0.017) ^a	0.011 (0.013)	0.165 (0.017) ^a	- 0.002 (0.013)
Industry 8's R&D	0.206 (0.027) ^a	0.076 (0.024) ^a	0.196 (0.027) ^a	0.065 (0.026) ^b
Industry 9's R&D	0.086 (0.015) ^a	0.005 (0.009)	0.085 (0.014) ^a	0.004 (0.009)
Industry 10's R&D	- 0.029 (0.026)	- 0.054 (0.023) ^a	- 0.035 (0.026)	- 0.056 (0.025) ^a
Foreign R&D	0.285 (0.007) ^a	- 0.095 (0.025) ^a	0.277 (0.013) ^a	- 0.088 (0.028) ^a
N	9,805		10,176	

For OPW specification, capital-labor ratio and own-industry domestic R&D of all industries are considered endogenous and are represented by their one period lagged values. The inter-industry R&D is considered as exogenous. Foreign R&D is treated as exogenous. In all regressions, country-, industry-, and time fixed effects are included.

a = significant at the 1% level, b = significant at the 5% level.

Table A12. Technology diffusion matrix using Olley-Pakes-Wooldridge method

Control variables	Receiving Industry	Source Industry			
		Pharmac eutical (industry 4)	Chemical s (industry 8)	Motor vehicles (industry 6)	Office accounting (industry 2)
	1. Aircraft and spacecraft	0.283 ^a	0.182 ^a	0.118 ^a	0.051 ^a
	2. Office, accounting and computing machinery	0.122 ^a	-0.051 ^a	-0.039 ^a	0.264^a
	3. Radio, TV and communication equip	0.087 ^a	0.017	-0.017	0.066 ^a
	4. Pharmaceuticals	0.150^a	0.105 ^a	0.021 ^b	0.019 ^b
	5. Medical, precision and optical instruments	0.171 ^a	0.103 ^a	0.037 ^a	-0.001
	6. Motor vehicles, trailers and semi-trailers	0.194 ^a	0.112 ^a	0.073^a	0.002
	7. Electrical machinery and apparatus, nec	0.145 ^a	0.106 ^a	0.022 ^b	0.015
	8. Chemicals excluding pharmaceuticals	0.226 ^a	0.162^a	0.069 ^a	0.044 ^a
	9. Railroad equip. and transport equip nec.	0.075 ^a	-0.029	0.005	-0.044 ^a
	10. Machinery and equipment, nec.	0.164 ^a	0.114 ^a	0.051 ^a	-0.009
	11. Coke, refined petroleum and nuclear fuel	0.224 ^a	0.150 ^a	0.095 ^a	0.027 ^a
	12. Rubber and plastics products	0.074 ^a	0.032	-0.004	-0.032 ^a
	13. Non-ferrous metals	0.112 ^a	0.054 ^b	0.009	0.011
	14. Iron and steel	0.145 ^a	0.102 ^a	0.036 ^a	0.018 ^b
	15. Building and repairing of ships and boats	0.098 ^a	0.072 ^a	0.003	0.020 ^b
	16. Other non-metallic mineral products	0.099 ^a	0.064 ^b	0.011	-0.015
	17. Fabricated metal products	0.093 ^a	0.069 ^a	0.015	-0.030 ^a
	18. Food products, beverages and tobacco	0.123 ^a	0.080 ^a	0.024 ^b	-0.014
	19. Manufacturing nec; recycling	0.047 ^b	0.037	-0.012	-0.037 ^a
	20. Pulp, paper, printing and publishing	0.100 ^a	0.048	0.008	-0.027 ^a
	21. Textiles, textile products, leather	0.137 ^a	0.095 ^a	0.045 ^a	-0.010
	22. Wood and products of wood and cork	0.134 ^a	0.082 ^a	0.015	0.011
	23. Computer, R&D and other business	0.047	0.168 ^a	0.123 ^a	0.097 ^b
	24. Post and telecommunications	0.047	0.168 ^a	0.123 ^a	0.097 ^b
	25. Transport and storage	0.047	0.168 ^a	0.123 ^a	0.097 ^b
	26. Other services	0.047	0.168 ^a	0.123 ^a	0.097 ^b
	27. Electricity, gas and water supply	0.117 ^a	0.073 ^a	0.028 ^b	-0.022
	28. Construction	0.017	0.008	-0.023	-0.095 ^a
	Industry 1's R&D	-0.052 ^a	-0.049 ^a	-0.050 ^a	-0.057 ^a
	Industry 2's R&D	0.010	0.011	0.010	-
	Industry 3's R&D	0.047 ^a	0.047 ^a	0.048 ^a	0.052 ^a
	Industry 4's R&D	-	0.137 ^a	0.131 ^a	0.121 ^a
	Industry 5's R&D	-0.000	0.001	0.002	-0.003
	Industry 6's R&D	0.025 ^a	0.027 ^a	-	0.022 ^a
	Industry 7's R&D	-0.009	-0.007	-0.010	-0.020 ^a
	Industry 8's R&D	0.084 ^a	-	0.081 ^a	0.108 ^a
	Industry 9's R&D	0.003	0.009	0.006	-0.007
	Industry 10's R&D	-0.076 ^a	-0.079 ^a	-0.078 ^a	-0.066 ^a
	Labor	0.540 ^a	0.557 ^a	0.550 ^a	0.575 ^a
	Capital	0.297 ^a	0.370 ^a	0.330 ^a	0.087 ^a
	18 industries R&D	0.076 ^a	0.062 ^a	0.064 ^a	0.071 ^a
	Foreign R&D	0.307 ^a	0.289 ^a	0.279 ^a	0.276 ^a
	N	9,856	9,856	9,856	9,856

Only labor is considered endogenous and is represented by its one period lagged value.

In all regressions, country-, industry-, and time fixed effects are included.

Standard errors are not reported.

a = significant at the 1% level, b = significant at the 5% level.

Appendix B: Data Note

The paper uses several OECD databases, with supplement data from other sources. The main data we have used from the OECD are: the Analytical Business Expenditure in Research and Development (ANBERD) database, Structural Analysis (STAN), Commodity, Trade and Production (COMTAP), Bilateral Trade (BTD) and input-output tables. We have also used data from the Groningen Growth and Development Centre (GGDC). As well, we have used data directly from the websites of national statistical agencies in Canada, Japan, the UK and the US. The industries for the study are based on International Standard Industrial Classification (ISIC) Revision 3 code and have used data from both old systems (ISIC Rev. 2), and Rev. 3. In what follows, we will provide a detailed data description and related concordance that we have used.

ANBERD: These data are available in two series: ANBERD 2 and ANBERD 3, the former based on ISIC Rev. 2, and the latter based on ISIC Rev. 3 industry code. ANBERD 2 data start from 1973 and covers till about 1995-97, and ANBERD 3 data start from 1987 and go at least till 2002 for most of the countries (except for Ireland in which case they stop at 2001). Belgium and Korea have no data on ANBERD 2, so for these two countries R&D data start from 1987. We have combined both data series. Industry value added deflator based on GGDC and STAN databases (more on these two databases later) are used to convert ANBERD current value data into 1995 prices and finally convert them into 1995 purchasing power parity (PPP) US dollar.

COMTAP and BTD: Trade data come from three OECD databases: (1) COMTAP for years 1970-1979, (2) BTD 2 for years 1980-1989 and (3) BTD 3 for years 1990-2003. Trade (import) data are complete for all 17 countries except for Korea, whose imports from sample countries are available only after 1994. However, data on exports by other sample countries to Korea are available since 1980. To combine the data for Korea's imports, we use the data on exports by sample countries to Korea as Korea's imports from those countries for year 1980 to 1993, and use Korea's imports data from 1994. For Germany, the import data are taken for West Germany from 1970 to 1989 and for united Germany from 1991 and onward. Since there are no bilateral trade data for services we use the average of manufacturing trade share as proxy for services industries. The trade data, which are in US dollar, are converted into 1995 purchasing power parity (PPP) dollar.

STAN: We use value added, gross fixed capital formation (investment), employment (persons engaged) and labor compensation data from the STAN database. We use mainly the STAN 3 database, which is based on ISIC Rev. 3 industry code, and start coverage from 1970 till more recent years. However, it appears that for early years for some countries, data that are available on STAN 2 are not available in STAN 3. In such cases, we use the STAN 2 database to cover the missing values in STAN 3. Furthermore, for some of the cells, which are empty in both STAN 2 and STAN 3 databases, we use data from an old CD to refill.

In STAN 3, value added data are available in nominal terms, and in real terms, i.e. as volumes. The former are in national currencies; the volumes are expressed as index numbers with national reference year equal to 100. Since, for the study, we need value added in value not as index, we converted the index into value. Second, since different countries index values are based on different reference year, we re-based the reference year for all countries in 1995 and converted them into 1995 PPP. By doing so, we have data on value added (both at constant and current prices) investment at current prices, employment, and compensation for the whole sample period.

GGDC: We have taken data for value added (both in current price and in constant price) and employment from GGDC. This dataset is comparable with the OECD STAN database but provides a dataset without gaps by complementing STAN with information from industry and services statistics and additional (historical) national accounts data for individual countries. The GGDC database have total of 57 industries and can be easily concorded into our sample 28 industries, except for two industries in which case we have to decompose these two GGDC industries into two each. In GGDC, the ISIC 24 is not split into ISIC 24x2423 and 2423. Similarly, industry ISIC 27 is not disaggregated into ISIC 271 and ISIC 272. To split GGDC 24 and 27 into two industries each, we used the value added at current price data from the STAN database where data on ISIC 24x2423, ISIC 2423, ISIC 271 and ISIC 272 are reported separately. We

computed the annual share of value added of 24x2423 and 2423 in ISIC 24 and used that share to decompose GGDC ISIC 24 into two categories. We did the same for ISIC 271 and 272, using the value added shares of these two industries in ISIC 27.

We combined the data on value added (both at current and constant price) and employment from the STAN and GGDC, taking data from the STAN for years 1973 to 1978 and from the GGDC for years 1979 and onward. Then using the combined value added data in current price and constant price of 1995, we calculated value added deflator, which was used to deflate the investment and R&D data. Then, we converted these national currency value added and investment data into 1995 PPP.

Supplementing and Estimating Missing Data

The variables that are used for the study are trade, value added, employment, R&D, labor compensation, and physical investment. The data on trade are almost complete except for a few years for Korea, so we have not estimated the missing values for this variable. The value added and employment data after 1979, when we had them from the GGDC, are complete. However, there are some missing values prior to 1979 for these variables. Among the three remaining variables, even though there are some data missing for R&D and labor compensation, the frequency of missing cells is more frequent in investment data. Below, we describe how we estimated some of the missing cells in investment data. We have also estimated a few missing cells in value added, employment, R&D expenditure and labor compensation using similar techniques.

The investment data are available in both current and constant prices. For the study, the preference would be to use constant price investment data, as they are based on more appropriate deflators. However, if we rely on constant price investment there will be a lot of missing cells. In terms of availability, the constant price investment data are a subset of current price investment in a sense that almost all data that are available in former series are also available in the latter but not vice versa. Hence, in this study, we have used the current price investment data and deflated them by value added deflators.

For the missing value, when possible, first we used national statistical agencies to refill the data if possible. This was done only for the UK, Canada, the US and Japan. For the UK, all 2001 and 2002 data were obtained from National Statistics UK. Data for Canada for most of 2001 and 2002 were taken from Statistics Canada. Similarly most of the data for 2001 and 2002 for the US were supplemented using Bureau of Economic Analysis "historical-cost investment in private fixed assets by industry" from the website. For Japan, the STAN database has investment data only in current price that too only till 1993. We supplemented these data by acquiring a file from Department of National Accounts, Economic and Social Research Institute in Japan, which has data from 1980 onward in constant price. Hence our investment series for Japan will be a mixed of two series: till 1979 we use the investment in current price by deflating with value added deflator, and from 1980 to 2002 we use data series which were already in constant price (using investment deflator).

A closer look at the industry level current price investment data shows that with sample period of 30 years, 28 industries and 17 countries ($30 \times 28 \times 17$), and total of 14,280 cells of information, 2,530 (about 17 percent) cells of investment data were missing. To estimate part of the missing cells we used three different approaches. The first approach is based on the assumption that the investment share of 3- or 4-digit level industry in 2-digit level industry remained the same as it was in the preceding three years. This method is used to estimate data mostly for industries at 3- and 4-digit level and for more recent years, 1999 through 2000. Since there are two 4-digit industries (ISIC 24x2423 and ISIC 2423) and three 3-digit industries (ISIC 351, ISIC 353, ISIC 352+359), we have used this method mostly for these five industries. The method, called Method 1, is given by the following equation:

$$(B1) \quad \hat{i}_i(3/4) \equiv \omega \hat{i}_i(2).$$

where $i_t(2)$ is the investment at 2-digit industry, $i_t(3/4)$ is the investment at 3 or 4 digit industries within that 2-digit industry in time period t ; $\bar{\omega} = \left(\sum_k i_{t-k}(3/4) / \sum_k i_{t-k}(2) \right) / 3$, ($k = 1, 2, 3$), is the average share of investment at 3- or 4-digit industry within its 2-digit industry investment in the preceding three years.

The second estimation method—Method 2—is used for those industries which have data available for at least three-fifths and less than four-fifths of the sample period (between 18 and 23 years). We used the change in current price value added to estimate investment as given below:

$$(B2) \quad i_{t+1} = i_t \exp \left[\ln(y_{t+1}/y_t) \right],$$

where i_t is investment in current price, and y is value added in current price. In most cases the data were available for early periods and we used (B2) to estimate data for later periods. In few cases, the data were available for later periods and were missing for earlier periods. In this case, we used

$i_{t-1} = i_t \exp \left[\ln(y_{t-1}/y_t) \right]$ to estimate the missing values. In very few cases, the data were missing in both ends with data available only for the middle period. In that case, we used (B2) to estimate data only for the later period and left the earlier period empty.

For those industries which have data for at least 24 years, we used the growth rate of investment—Method 3—to estimate investment for current years as follows:

$$(B3) \quad i_{t+1} = i_t \left[1 + \ln(i_t/i_{t-1}) \right]$$

Equation (A3) is good to estimate data for later period given that the earlier period data were available. To refill data for earlier period, we used $i_{t-1} = i_t \left[1 + \ln(i_t/i_{t-1}) \right]$.

For investment, 140 cells were filled up using method 1; about 816 cells were filled up using Method 2, and about 196 cells were filled up using Method 3. Hence, altogether 1,152 of the 2,530 missing cells were filled up. The remaining 1,378 were left empty either because the data for industries were empty throughout or were available for less than 18 years, the cut off number of years for data refinement. Then we used value added deflators to convert adjusted current price investment into constant price, and further converted them into 1995 PPP dollar.

In very few cases, we have augmented the value added and deflator prior to 1979 using Method (3). In case of current value added, we filled 274 cells and in case of deflator, we filled 460 cells. In the case of R&D, we filled 430 cells using this method.

